Baskar
09/853079

Claim 3
Sept IDS
general search
for B
and S

FILE 'REGISTRY' ENTERED AT 12:05:21 ON 30 AUG 2002
L1 918 S [EG][AT][GV][WG][PS]S/SQSP

FILE 'HCAPLUS' ENTERED AT 12:06:28 ON 30 AUG 2002
L2 491 S L1
L3 7 S L2 AND MICROTI

L3 ANSWER 1 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:833517 HCAPLUS
DOCUMENT NUMBER: 135:367756
TITLE: Babesia **microti** antigens and methods for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond L.; Sleath, Paul R.; McNeill, Patricia D.; Homer, Mary J.; Secrist, Heather
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: PCT Int. Appl., 195 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001085947	A2	20011115	WO 2001-US15192	20010509
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 2000-569098	A	20000510
US 2000-605724	A	20000627
US 2000-656688	A	20000907
US 2000-685436	A	20001010
US 2000-737178	A	20001213
US 2001-794764	A	20010226
US 1996-723142	A2	19961001
US 1997-845258	A2	19970424
US 1997-990571	A2	19971211
WO 1998-US26437	A2	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405

AB Compds. and methods for the diagnosis and treatment of B. **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and immunogenic

compns. comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

IT **206205-11-0**, Antigen BMNI-1 (Babesia **microti**)
206205-12-1, Antigen BMNI-2 (Babesia **microti**)
206205-13-2, Antigen BMNI-3 (Babesia **microti**)
206205-16-5, Antigen BMNI-6 (Babesia **microti**)
206205-20-1, Antigen BMNI-12 (Babesia **microti**)
206205-21-2, Antigen BMNI-13 (Babesia **microti**)
206205-23-4, Antigen BMNI-16 (Babesia **microti**)
RL: BUU (Biological use, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amino acid sequence; of Babesia **microti** antigens used in diagnosis and treatment of Babesia **microti** infection)

IT **206205-33-6 206205-35-8 206205-36-9**
227296-22-2 227296-23-3 227296-26-6
227296-30-2 227296-31-3 227296-32-4
227296-33-5 227296-34-6 227296-35-7,
Antigen MN2 (Bombesia microtia fragment) **227296-36-8**
227296-37-9, Antigen MN3 (Bombesia microtia fragment)
RL: PRP (Properties)
(unclaimed protein sequence; babesia **microti** antigens and methods for the diagnosis and treatment of Babesia **microti** infection)

IT **205488-48-8 205488-54-6 334074-87-2**
334074-88-3 334074-89-4 334074-90-7
RL: PRP (Properties)
(unclaimed sequence; babesia **microti** antigens and methods for the diagnosis and treatment of Babesia **microti** infection)

L3 ANSWER 2 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:748300 HCAPLUS
DOCUMENT NUMBER: 135:299589
TITLE: Nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond L.; Sleath, Paul R.; McNeill, Patricia D.
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 26 pp., Cont.-in-part of U. S. Ser. No. 685,436.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2001029295	A1	20011011	US 2000-737178	20001213
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424
US 6214971	B1	20010410	US 1997-990571	19971211

WO 9929869 A1 19990617 WO 1998-US26437 19981211
W: AU, CA, JP, MX, NZ
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
WO 2000060090 A1 20001012 WO 2000-US9136 20000405
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
WO 2001085947 A2 20011115 WO 2001-US15192 20010509
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.:
US 1996-723142 A2 19961001
US 1997-845258 A2 19970424
US 1997-990571 A2 19971211
WO 1998-US26437 A2 19981211
US 1999-286488 A2 19990405
US 2000-528784 A2 20000317
WO 2000-US9136 A2 20000405
US 2000-569098 A2 20000510
US 2000-605724 A2 20000627
US 2000-656688 A2 20000907
US 2000-685436 A2 20001010
US 2000-737178 A 20001213
US 2001-794764 A 20010226

AB Compds. and methods for the diagnosis and treatment of B. microtiinfection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. microtiantigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and immunogenic compns. comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. microtiinfection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

IT **206205-36-9**
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection)

IT **227296-26-6 227296-30-2 227296-31-3 227296-32-4 227296-34-6 227296-35-7,**
Antigen MN2 (Bombesia microtia fragment) **227296-36-8**

227296-37-9, Antigen MN3 (Bombesia microtia fragment)
334074-87-2 334074-88-3 334074-89-4
334074-90-7
RL: PRP (Properties)
(unclaimed protein sequence; nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection)

L3 ANSWER 3 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:255942 HCAPLUS
DOCUMENT NUMBER: 134:294507
TITLE: Compounds and methods for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: U.S., 78 pp., Cont.-in-part of U.S. Ser. No. 845,258.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6214971	B1	20010410	US 1997-990571	19971211
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424
WO 9929869	A1	19990617	WO 1998-US26437	19981211
W: AU, CA, JP, MX, NZ				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
AU 9918204	A1	19990628	AU 1999-18204	19981211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 1996-723142	A2	19961001
US 1997-845258	A2	19970424
US 1997-990571	A	19971211
WO 1998-US26437	W	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405
US 2000-569098	A2	20000510
US 2000-605724	A2	20000627
US 2000-656688	A2	20000907
US 2000-685436	A2	20001010

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.
IT **206205-11-0 206205-12-1 206205-13-2**

206205-16-5 206205-20-1 206205-21-2
206205-23-4
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; compds. and methods for the diagnosis and treatment of Babesia **microti** infection)

IT **206205-33-6 206205-35-8 206205-36-9**
227296-22-2 227296-23-3 227296-26-6
227296-30-2 227296-31-3 227296-32-4
227296-33-5 227296-34-6 227296-35-7,
Antigen MN2 (Bombesia microtia fragment) **227296-36-8**
227296-37-9, Antigen MN3 (Bombesia microtia fragment)
334074-87-2
RL: PRP (Properties)
(unclaimed protein sequence; compds. and methods for the diagnosis and treatment of Babesia **microti** infection)

IT **205488-48-8 205488-54-6 334074-88-3**
334074-89-4 334074-90-7
RL: PRP (Properties)
(unclaimed sequence; compds. and methods for the diagnosis and treatment of Babesia **microti** infection)

REFERENCE COUNT: 13 THERE ARE 13 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 4 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:91448 HCAPLUS
DOCUMENT NUMBER: 134:158493
TITLE: Nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond; Sleath, Paul R.
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: U.S., 66 pp., Cont.-in-part of U.S. Ser. No. 723,142.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6183976	B1	20010206	US 1997-845258	19970424
US 6306396	B1	20011023	US 1996-723142	19961001
EP 834567	A2	19980408	EP 1997-117067	19971001
EP 834567	A3	19990331		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6214971	B1	20010410	US 1997-990571	19971211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:
US 1996-723142 A2 19961001
US 1997-845258 A 19970424
US 1997-990571 A2 19971211
WO 1998-US26437 A2 19981211
US 1999-286488 A2 19990405
US 2000-528784 A2 20000317
WO 2000-US9136 A2 20000405

US 2000-569098 A2 20000510
US 2000-605724 A2 20000627
US 2000-656688 A2 20000907
US 2000-685436 A2 20001010

OTHER SOURCE(S): MARPAT 134:158493

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. The DNA sequences encoding B. **microti** antigens were prepd. by screening a B. **microti** expression library with sera obtained from patients infected with B. **microti**. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

IT **205488-48-8 205488-54-6 206205-11-0**
206205-12-1 206205-13-2 206205-16-5
206205-20-1 206205-21-2 206205-23-4
206205-33-6 206205-35-8 206205-36-9
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); OCCU (Occurrence); USES (Uses)
(amino acid sequence; nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection)

REFERENCE COUNT: 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 5 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2000:282570 HCAPLUS
DOCUMENT NUMBER: 133:72606
TITLE: Serological expression cloning of novel immunoreactive antigens of Babesia **microti**
AUTHOR(S): Lodes, Michael J.; Houghton, Raymond L.; Bruinsma, Elizabeth S.; Mohamath, Raodoh; Reynolds, Lisa D.; Benson, Darin R.; Krause, Peter J.; Reed, Steven G.; Persing, David H.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA
SOURCE: Infection and Immunity (2000), 68(5), 2783-2790
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Increased recognition of the prevalence of human babesiosis in the United States, together with rising concern about the potential for transmission of this infection by blood transfusion, has provided motivation to develop definitive serol. and mol. tests for the causative agent, Babesia **microti**. To develop more sensitive and specific assays for B. **microti**, the authors screened a genomic expression library with patient serum pools. This screening resulted in the identification of three classes of novel genes and an addnl. two novel, unrelated genes, which together

encode a total of 17 unique B. **microti** antigens. The first class (BMN1-2 family) of genes encodes seven closely related antigens with a degenerate six-amino-acid repeat that shows limited homol. to Plasmodium sp. merozoite and sporozoite surface antigens. A second class (BMN1-8 family) of genes encodes six related antigens, and the third class (BMN1-17 family) of genes encodes two related antigens. The two remaining genes code for novel and unrelated sequences. Among the three classes of antigens and remaining novel sequences, five were chosen to code for the most immunodominant antigens (BMN1-2, -9, -15, and -17 and MN-10). Western blot anal. with the resulting recombinant proteins indicated that these antigens were targets of humoral immune responses during B. **microti** infection in humans.

IT **278626-95-2 278626-96-3 278626-97-4**
278626-98-5 278626-99-6
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(amino acid sequence; serol. expression cloning of novel immunoreactive antigens of Babesia **microti**)

REFERENCE COUNT: 36 THERE ARE 36 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 6 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1999:388301 HCAPLUS
DOCUMENT NUMBER: 131:85409
TITLE: Antigen and gene sequences for diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond; Sleath, Paul R.; Persing, David; Bruinsma, Elizabeth
PATENT ASSIGNEE(S): Corixa Corporation, USA; Mayo Foundation for Medical Education and Research
SOURCE: PCT Int. Appl., 126 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9929869	A1	19990617	WO 1998-US26437	19981211
W: AU, CA, JP, MX, NZ				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 6214971	B1	20010410	US 1997-990571	19971211
AU 9918204	A1	19990628	AU 1999-18204	19981211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 1997-990571	A	19971211
US 1996-723142	A2	19961001
US 1997-845258	A2	19970424
WO 1998-US26437	W	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405
US 2000-569098	A2	20000510
US 2000-605724	A2	20000627

US 2000-656688 A2 20000907
US 2000-685436 A2 20001010

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Nine antigens share some homol., contain a degenerate repeat of six amino acids with 9-22 repeats occurring in each antigen, and bear some similarity to a Plasmodium falciparum merozoite surface antigen gene. A second group of five antigens bear some homol. to each other but do not show homol. to any previously identified sequences. Two synthetic peptides (BABS-1 and BABS-4) were made to the repeat region of isolated B. **microti** antigen BMNI-3. Twelve BMNI-6 homologs were obtained from hamster and human patients. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

IT **205488-48-8**
RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(BABS-1 peptide fragment of repeat region of BMNI-3; antigen and gene sequences for diagnosis and treatment of Babesia **microti** infection)

IT **205488-54-6**
RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(BABS-4 peptide fragment of repeat region of BMNI-3; antigen and gene sequences for diagnosis and treatment of Babesia **microti** infection)

IT **206205-11-0 206205-12-1 206205-13-2**
206205-16-5 206205-20-1 206205-21-2
206205-33-6 206205-35-8 227296-22-2
227296-23-3 227296-26-6 227296-30-2
227296-31-3 227296-32-4 227296-33-5
227296-34-6 227296-35-7, Antigen MN2 (Bombesia microtia fragment) **227296-36-8 227296-37-9**, Antigen MN3 (Bombesia microtia fragment)
RL: ANT (Analyte); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); OCCU (Occurrence); USES (Uses)
(amino acid sequence; antigen and gene sequences for diagnosis and treatment of Babesia **microti** infection)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 7 OF 7 HCAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1998:229027 HCAPLUS
DOCUMENT NUMBER: 128:292989
TITLE: Antigens of Babesia microtia and their use in the diagnosis and treatment of infection
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton,

Raymond; Sleath, Paul R.
PATENT ASSIGNEE(S): Corixa Corp., USA
SOURCE: Eur. Pat. Appl., 113 pp.
CODEN: EPXXDW
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 834567	A2	19980408	EP 1997-117067	19971001
EP 834567	A3	19990331		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424

PRIORITY APPLN. INFO.: US 1996-723142 A 19961001
US 1997-845258 A 19970424

AB Antigens and epitopes of Babesia **microti** that can be used in the diagnosis and treatment of infection are described. Genes for the antigens or antibodies against them can be used in the detection of B. **microti**. CDNAS for these antigens were cloned by screening an expression library in .lambda.ZAP with antiserum to B. **microti**.

IT **205488-48-8 205488-54-6 206205-11-0**
206205-12-1 206205-13-2 206205-16-5
206205-20-1 206205-21-2 206205-23-4
206205-33-6 206205-35-8 206205-36-9
RL: ANT (Analyte); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(amino acid sequence; antigens of Babesia microtia and their use in diagnosis and treatment of infection)

E1 THROUGH E32 ASSIGNED

FILE 'REGISTRY' ENTERED AT 12:07:45 ON 30 AUG 2002
L4 32 SEA FILE=REGISTRY ABB=ON PLU=ON (205488-48-8/BI OR 205488-54-6/BI OR 206205-11-0/BI OR 206205-12-1/BI OR 206205-13-2/BI OR 206205-16-5/BI OR 206205-20-1/BI OR 206205-21-2/BI OR 206205-33-6/BI OR 206205-35-8/BI OR 206205-36-9/BI OR 206205-23-4/BI OR 227296-26-6/BI OR 227296-30-2/BI OR 227296-31-3/BI OR 227296-32-4/BI OR 227296-34-6/BI OR 227296-35-7/BI OR 227296-36-8/BI OR 227296-37-9/BI OR 227296-22-2/BI OR 227296-23-3/BI OR 227296-33-5/BI OR 334074-87-2/BI OR 334074-88-3/BI OR 334074-89-4/BI OR 334074-90-7/BI OR 278626-95-2/BI OR 278626-96-3/BI OR 278626-97-4/BI OR 278626-98-5/BI OR 278626-99-6/BI)

L5 32 L4 AND L1

L5 ANSWER 1 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **334074-90-7** REGISTRY
CN 90: PN: US6214971 FIGURE: 6 unclaimed sequence (9CI) (CA INDEX NAME)
OTHER NAMES:

CN 10: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 204: PN: WO0185947 FIGURE: 6 unclaimed sequence
CI MAN
SQL 147

SEQ 1 YITLFLMSGA VFAGDTDREA GGPSGTVGPS EAGGPSEAGG PSEAGGPSEA
== ========== ========== ==========
51 GGPSEAGGPS EAGGPSHAGG PSEAGGPSGT GWPSEAGWPS EAGWPSEAGW
========== ====== ======== ========== ==========
101 PSEAGWPSEA GWPSERFGYQ LLWYSRRIVI FNEIYLSHIY EHSVMIL
========== ====
HITS AT: 19-66, 73-114

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

L5 ANSWER 2 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **334074-89-4** REGISTRY
CN 86: PN: US6214971 FIGURE: 6 unclaimed sequence (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 200: PN: WO0185947 FIGURE: 6 unclaimed sequence
CN 6: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CI MAN
SQL 138

SEQ 1 AGDTDREAGG PSGTVGPSSA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
==== ======== ====== ========== ==========
51 GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA GGPSEAGWPS EAGWPSEAGG
========== ========== ========== ========== ==========
101 PSGTGWPSEA GWPSEAGWPS EAGWPSEAGW PSERFGYQ
========== ========== ========== ==
HITS AT: 7-18, 25-132

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

L5 ANSWER 3 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **334074-88-3** REGISTRY
CN 81: PN: US6214971 FIGURE: 6 unclaimed sequence (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 195: PN: WO0185947 FIGURE: 6 unclaimed sequence
CN 1: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CI MAN
SQL 111

SEQ 1 GDTDREAGGP SGTVGPSEAG GPSEAGGPSG TVGPSEAGGP SEAGGPSGTG
===== ========== ========== ========== ==========
51 WPSEAGGPSG TVGPSEAGGP SEAGGPSGTG WPSGTGWPSE VGWPSERFGY
========== ========== ========== =========
101 QLLWYSRRIV I

HITS AT: 6-89

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

L5 ANSWER 4 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **334074-87-2** REGISTRY
CN 80: PN: US6214971 FIGURE: 6 unclaimed protein (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 194: PN: WO0185947 FIGURE: 6 unclaimed sequence
CN 1: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CI MAN
SQL 112

SEQ 1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS GTVGPSEAGG PSEAGGPSGT
==== ========== ========== ========== ==========
51 GWPSEAGGPS GTVGPSEAGG PSEAGGPSGT GWPSGTGWPS EVGWPSERFG
========== ========== ========== ==========
101 YQLLWYSRRI VI
HITS AT: 7-90

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

L5 ANSWER 5 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **278626-99-6** REGISTRY
CN Antigen BMN1-13 (Babesia microtia strain MN1 clone bmn1-13 precursor) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF206524-derived protein GI 7716011
CI MAN
SQL 262

SEQ 1 MVSFKSILVP YITLFLMSGA VFASDTDPEA GGPSGTVGPS EAGGPSEAGG
== ========== ==========
51 PSGTGWPSEA GGPSEAGGPS GTGWPSEAGW SSERFGYQLL PYSRRIVTFN
========== ========== ========== ==
101 EVCLSYIYKH SVMILERDRV NDGHKDYIEE KTKEKNKLKK ELEKCFPEQY
151 SLMKKEELAR IFDNASTISS KYKLLVDEIS NKAYGTLEGP AADNFDHFRN
201 IWKSIVLKDM FIYCDLLLQH LIYKFYYDNT INDIKKNFDE SKSKALVLRD
251 KITKKDVYVN DH
HITS AT: 29-82

REFERENCE 1: 133:72606

L5 ANSWER 6 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **278626-98-5** REGISTRY
CN Antigen BMN1-7 (Babesia microtia strain MN1 clone bmn1-7 precursor) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF206250-derived protein GI 7715998
CI MAN

```
SQL  289

SEQ      1 MVSFKSILVP YITLFLMSGA VFAGDTDREA GGPSGTVGPS EAGGPSEAGG
                                     == ========== ==========
        51 PSEAGGPSEA GGPSEAGGPS EAGGPSEAGG PSGTVGPSEA GGPSEAGGPS
           ========== ========== ========== ========== ==========
       101 EAGGPSEAGW PSEAGWPSEA GWPSEAGWPS EAGWPSEAGW PSERFGYQLL
           ========== ========== ========== ========== ==
       151 WYSRRIVIFN EIYLSHIYEH SVMILERDRV NDGHKDYIEE KTKEKNKLKK
       201 ELEKCFPEQY SLMKKEELAR IIDNASTISS KYKLLVDEIS NKAYGTLEGP
       251 AADDFDHFRN IWKSIVPKNM FLYCDLLLKH LIRLTPRKS
HITS AT:   29-142

REFERENCE   1:  133:72606

L5   ANSWER 7 OF 32  REGISTRY  COPYRIGHT 2002 ACS
RN   278626-97-4  REGISTRY
CN   Antigen BMN1-6 (Babesia microtia strain MN1 clone bmn1-6 precursor)
     (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AF206249-derived protein GI 7715996
CI   MAN
SQL  298

SEQ      1 MVSFKSILVP YITLFLMSGA VFAGDTDREA GGPSGTVGPS EAGGPSEAGG
                                     == ========== ==========
        51 PSEAGGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSGT GWPSEAGWPS
           ========== ========== ========== ========== ==========
       101 EAGWPSEAGW PSEAGWPSEA GWPSERFGYQ LLWYSRRIVI FNEIYLSHIY
           ========== ========== ====
       151 EHSVMILERD RVNDGHKDYI EEKTKEKNKL KKELEKCFPE QYSLMKKEEL
       201 ARIIDNASTI SSKYKLLVDE ISNKAYGTLE GPAADDFDHF RNIWKSIVPK
       251 NMFLYCDLLL KHLIRKFYCD NTINDIKKNF DDIEKLGCFQ ARSFLPVN
HITS AT:   29-124

REFERENCE   1:  133:72606

L5   ANSWER 8 OF 32  REGISTRY  COPYRIGHT 2002 ACS
RN   278626-96-3  REGISTRY
CN   Antigen BMN1-3 (Babesia microtia strain MN1 clone bmn1-3 precursor)
     (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AF206245-derived protein GI 7715986
CI   MAN
SQL  362

SEQ      1 MVSFKSILVP YITLFLMSGA VFASDTDPEA GGPSEAGGPS GTVGPSEAGG
                                     == ========== ==========
        51 PSEAGGPSGT GWPSEAGGPS EAGGPSEAGG PSEAGGPSGT GWPSGTGWPS
           ========== ========== ========== ========== ==========
       101 EAGWSSERFG YQLLPYSRRI VIFNEVCLSY IYKHSVMILE RDRVNDGHKD
           ======
       151 YIEEKTKEKN KLKKELEKCF PEQYSLMKKE ELARIFDNAS TISSKYKLLV
       201 DEISNKAYGT LEGPAADNFD HFRNIWKSIV LKDMFIYCDL LLQHLIYKFY
       251 YDNTVNDIKK NFDESKSKAL VLRDKITKKD GDYNTHFEDM IKELNSAAEE
       301 FNKIVDIMIS NIGDYDEYDS IASFKPFLSM ITEITKITKV SNVIIPGIKA
       351 LTLTVFLIFI TK
HITS AT:   29-106
```

REFERENCE 1: 133:72606

L5 ANSWER 9 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **278626-95-2** REGISTRY
CN Antigen BMN1-2 (Babesia microtia strain MN1 clone bmn1-2 precursor) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF206244-derived protein GI 7715984
CI MAN
SQL 326

```
SEQ      1 MVSFKSILVP YITLFLMSGA VFASDTDPEA GGPSEAGGPS GTVGPSEAGG
                                           == ========== ==========
        51 PSEAGGPSGT VGPSEAGGPS EAGGPSGTGW PSEAGGPSEA GGPSGTVGPS
           ========== ========== ========== ========== ==========
       101 EAGGPSEAGG PSGTGWPSEA GGPSEAGGPS EAGGPSEAGG PSGTGWPSGT
           ========== ========== ========== ========== ==========
       151 GWPSEAGWSS ERFGYQLLPY SRRIVIFNEV CLSYIYKHSV MILERDRVND
           ==========
       201 GHKDYIEEKT KEKNKLKKEL EKCFPEQYSL MKKEELARIF DNASTISSKY
       251 KLLVDEISNK AYGTLEGPAA DNFDHFRNIW KSIVLKDMFI YCDLLLQHLI
       301 YKFYYDNTVN DIKKNFDESW TQTLKE
HITS AT:   29-160
```

REFERENCE 1: 133:72606

L5 ANSWER 10 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-37-9** REGISTRY
CN Antigen MN3 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 11: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 47: PN: WO0185947 SEQID: 78 unclaimed protein
CN 74: PN: WO0060090 SEQID: 78 unclaimed protein
CN 78: PN: US6214971 SEQID: 78 unclaimed protein
CN 79: PN: US6214971 SEQID: 79 unclaimed protein
CN 91: PN: US6214971 FIGURE: 6 unclaimed sequence
CN 92: PN: US6214971 FIGURE: 6 unclaimed sequence
CN Antigen MRT (Bombesia microtia fragment)
CI MAN
SQL 120

```
SEQ      1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
               ==== ========== ========== ========== ==========
        51 GGPSEAGGPS EAGGPSEAGG PSGTGWPSEA GWPSEAGWPS EAGWPSEAGW
           ========== ========== ========== ========== ==========
       101 PSEAGWPSER FGYQLLWYSR
           =========
HITS AT:   7-108
```

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 11 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-36-8** REGISTRY
CN Antigen MN1PAT (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 46: PN: WO0185947 SEQID: 77 unclaimed protein
CN 73: PN: WO0060090 SEQID: 77 unclaimed protein
CN 77: PN: US6214971 SEQID: 77 unclaimed protein
CN 89: PN: US6214971 FIGURE: 6 unclaimed sequence
CN 9: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CI MAN
SQL 113

SEQ 1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
==== ========== ========== ========== ==========
51 GGPSEAGGPS EAGGPSGTGW PSEAGWPSEA GWPSEAGWPS EAGWPSEAGW
========== ========== ========== ========== ==========
101 PSERFGYQLL WYS
==
HITS AT: 7-102

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 12 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-35-7** REGISTRY
CN Antigen MN2 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 45: PN: WO0185947 SEQID: 76 unclaimed protein
CN 72: PN: WO0060090 SEQID: 76 unclaimed protein
CN 76: PN: US6214971 SEQID: 76 unclaimed protein
CN 88: PN: US6214971 FIGURE: 6 unclaimed sequence
CN 8: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CI MAN
SQL 94

SEQ 1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
==== ========== ========== ========== ==========
51 GGPSEAGGPS GTGWPSEAGW PSEAGWPSEA GWPSEAGWPS EAGW
========== ========== ========== ==========
HITS AT: 7-90

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 13 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-34-6** REGISTRY
CN Antigen MN1HAM (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 44: PN: WO0185947 SEQID: 75 unclaimed protein
CN 71: PN: WO0060090 SEQID: 75 unclaimed protein
CN 75: PN: US6214971 SEQID: 75 unclaimed protein
CN 7: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 87: PN: US6214971 FIGURE: 6 unclaimed sequence
CI MAN
SQL 118

```
SEQ      1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
             ==== ========== ========== ========== ==========
        51 GGPSEAGGPS EAGGPSGTGW PSEAGWPSEA GWPSEAGWPS EAGWPSEAGW
           ========== ========== ========== ========== ==========
       101 PSERFGYQLL WYSRRIVI
           ==
```

HITS AT: 7-102

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 14 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-33-5** REGISTRY
CN Antigen RIFS (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 43: PN: WO0185947 SEQID: 74 unclaimed protein
CN 70: PN: WO0060090 SEQID: 74 unclaimed protein
CN 74: PN: US6214971 SEQID: 74 unclaimed protein
CI MAN
SQL 138

```
SEQ      1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
             ==== ========== ========== ========== ==========
        51 GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA GGPSEAGWPS EAGWPSEAGG
           ========== ========== ========== ========== ==========
       101 PSGTGWPSEA GWPSEAGWPS EAGWPSEAGW PSERFGYQ
           ========== ========== ========== ==
```

HITS AT: 7-132

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 133:295360

REFERENCE 4: 131:85409

L5 ANSWER 15 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-32-4** REGISTRY
CN Antigen BI2018 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 42: PN: WO0185947 SEQID: 73 unclaimed protein
CN 5: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 69: PN: WO0060090 SEQID: 73 unclaimed protein
CN 73: PN: US6214971 SEQID: 73 unclaimed protein
CN 85: PN: US6214971 FIGURE: 6 unclaimed sequence
CI MAN
SQL 136

SEQ 1 GDTDREAGGP SGTVGPSEAG GPSEAGGPSE AGGPSEAGGP SEAGGPSEAG
51 GPSEAGGPSE AGGPSEAGGP SEAGGPSEAG WPSEAGWPSE AGGPSGTGWP
101 SEAGWPSEAG WPSEAGWPSE AGWPSERFGY QLLWYS

HITS AT: 6-125

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 16 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-31-3** REGISTRY
CN Antigen BI2253 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 41: PN: WO0185947 SEQID: 72 unclaimed protein
CN 4: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 68: PN: WO0060090 SEQID: 72 unclaimed protein
CN 72: PN: US6214971 SEQID: 72 unclaimed protein
CN 84: PN: US6214971 FIGURE: 6 unclaimed sequence
CI MAN
SQL 116

SEQ 1 EAGGPSGTVG PSEAGGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
51 GGPSEAGGPS EAGGPSEAGW PSEAGWPSEA GGPSGTGWPS EAGWPSEAGW
101 PSEAGWPSEA GWPSER

HITS AT: 1-114

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 17 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-30-2** REGISTRY
CN Antigen BI2259 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 3: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 40: PN: WO0185947 SEQID: 71 unclaimed protein
CN 71: PN: US6214971 SEQID: 71 unclaimed protein
CN 83: PN: US6214971 FIGURE: 6 unclaimed sequence
CI MAN
SQL 136

```
SEQ      1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
             ==== ========== ========== ========== ==========
        51 GGPSEAGGPS EAGGPSEAGG PSEAWGPSEA GWPSEAGGPS GTGWPSEAGW
           ========== ========== ========== ========== ==========
       101 PSEAGWPSEA GWPSEAGWPS ERFGYQLLWY SRRIVI
           ========== ==========
```

HITS AT: 7-120

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 18 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-26-6** REGISTRY
CN Antigen BI2227 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 2: PN: US20010029295 FIGURE: 6A-6B unclaimed protein
CN 39: PN: WO0185947 SEQID: 70 unclaimed protein
CN 67: PN: WO0060090 SEQID: 70 unclaimed protein
CN 70: PN: US6214971 SEQID: 70 unclaimed protein
CN 82: PN: US6214971 FIGURE: 6 unclaimed sequence
CI MAN
SQL 118

```
SEQ      1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS EAGGPSEAGG PSEAGGPSEA
             ==== ========== ========== ========== ==========
        51 GGPSEAGGPS EAGGPSEAGG PSEAWGPSEA GWPSEAGGPS GTGWPSEAGW
           ========== ========== ========== ========== ==========
       101 PSEAGWPSEA GWPSEAGW
           ========== ====
```

HITS AT: 7-114

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

L5 ANSWER 19 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-23-3** REGISTRY
CN Antigen BI1053 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 38: PN: WO0185947 SEQID: 69 unclaimed protein
CN 66: PN: WO0060090 SEQID: 69 unclaimed protein
CN 69: PN: US6214971 SEQID: 69 unclaimed protein
CI MAN
SQL 105

SEQ 1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS GTVGPSEAGG PSEAGGPSGT
51 GWPSEAGGPS GTVGPSEAGG PSEAGGPSGT GWPSGTGWPS EVGWPNEPFG
101 YHLLW
HITS AT: 7-90

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 133:295360

REFERENCE 4: 131:85409

L5 ANSWER 20 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **227296-22-2** REGISTRY
CN Antigen BI254 (Bombesia microtia fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 37: PN: WO0185947 SEQID: 68 unclaimed protein
CN 65: PN: WO0060090 SEQID: 68 unclaimed protein
CN 68: PN: US6214971 SEQID: 68 unclaimed protein
CI MAN
SQL 101

SEQ 1 AGDTDREAGG PSGTVGPSEA GGPSEAGGPS GTVGPSEAGG PSEAGGPSGT
51 GWPSEAGGPS GTVGPSEAGG PSEAGGPSGT GWPSGTGWPS EVGWPIEPFG
101 Y
HITS AT: 7-90

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 133:295360

REFERENCE 4: 131:85409

L5 ANSWER 21 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-36-9** REGISTRY
CN Antigen (Babesia microtia 367-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 18: PN: WO0185947 SEQID: 49 unclaimed protein

CN 2: PN: US20010029295 SEQID: 49 claimed protein
CN 46: PN: WO0060090 SEQID: 49 unclaimed protein
CN 49: PN: US6183976 SEQID: 49 claimed protein
CN 49: PN: US6214971 SEQID: 49 unclaimed protein
CN Antigen BMNI-3 (Babesia microtia isoform 2 fragment)
CI MAN
SQL 367

SEQ 1 MVSFKSILVP YITLFLMSGA VFASDTDPEA GGPSEAGGPS GTVGPSEAGG
PSEAGGPSGT GWPSEAGGPS EAGGPSEAGG PSEAGGPSGT GSEAGGWPSG
TGWPSEAGWS SERFGYQLLP YSRRIVIFNE VCLSYIYKHS VMILERDRVN
DGHKDYIEEK TKEKNKLKKE LEKCFPEQYS LMKKEELARI FDNASTISSK
YKLLVDEISN KAYGTLEGPA ADNFDHFRNI WKSIVLKDMF IYCDLLLQHL
IYKFYYDNTV NDIKKNFDES KSKALVLRDK ITKKDGDYNT HFEDMIKELN
SAAEEFNKIV DIMISNIGDY DEYDSIASFK PFLSMITEIT KITKVSNVII
PGIKALTLTV FLIFITK
HITS AT: 29-88, 100-111

REFERENCE 1: 135:367756

REFERENCE 2: 135:299589

REFERENCE 3: 134:294507

REFERENCE 4: 134:158493

REFERENCE 5: 133:295360

REFERENCE 6: 128:292989

L5 ANSWER 22 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-35-8** REGISTRY
CN Antigen (Babesia microtia 294-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 14: PN: WO0185947 SEQID: 46 unclaimed protein
CN 45: PN: WO0060090 SEQID: 46 unclaimed protein
CN 46: PN: US6183976 SEQID: 46 claimed protein
CN 46: PN: US6214971 SEQID: 46 unclaimed protein
CN Antigen BMNI-7 (Babesia microtia isoform fragment)
CN Antigen BMNI-7 (Bombesia microtia reverse complement fragment)
CI MAN
SQL 294

SEQ 1 LWFIKMVSFK SILVPYITLF LMSGAVFAGD TDREAGGPSG TVGPSEAGGP
SEAGGPSEAG GPSEAGGPSE AGGPSEAGGP SEAGGPSGTV GPSEAGGPSE
AGGPSEAGGP SEAGWPSEAG WPSEAGWPSE AGWPSEAGWP SEAGWPSERF
GYQLLWYSRR IVIFNEIYLS HIYEHSVMIL ERDRVNDGHK DYIEEKTKEK
NKLKKELEKC FPEQYSLMKK EELARIIDNA STISSKYKLL VDEISNKAYG
TLEGPAADDF DHFRNIWKSI VPKNNFLYCD LLLKHLIRLT PRKS
HITS AT: 34-147

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 23 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-33-6** REGISTRY
CN Antigen (Babesia microtia 154-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 12: PN: WO0185947 SEQID: 44 unclaimed protein
CN 43: PN: WO0060090 SEQID: 44 unclaimed protein
CN 44: PN: US6183976 SEQID: 44 claimed protein
CN 44: PN: US6214971 SEQID: 44 unclaimed protein
CN Antigen BMNI-5 (Babesia microtia isoform 2 fragment)
CN Antigen BMNI-5 (Bombesia microtia reverse complement 154-amino acid fragment)
CI MAN
SQL 154

```
SEQ      1 LWFIKMVSFK SILVPYITLF LMSGAVFAGD TDREAGGPSG TVGPSEAGGP
                                              ======= ==========
        51 SEAGGPSEAG GPSEAGGPSE AGGPSEAGGP SEAGGPSEAG GPSGTGWPSE
           ========== ========== ========== ========== ==========
       101 AGGPSEAGGP SEAGGPSGTG WPSEAGWPSE AGWPSEAGWP SEAGWPSEAG
           ========== ========== ========== ========== ==========
       151 WPSE
           ===
```

HITS AT: 34-153

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 24 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-23-4** REGISTRY
CN Antigen (Babesia microtia 128-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 250: PN: WO0185947 SEQID: 31 claimed protein
CN 30: PN: WO0060090 SEQID: 31 unclaimed protein
CN 31: PN: US6183976 SEQID: 31 claimed protein
CN 31: PN: US6214971 SEQID: 31 claimed protein
CN Antigen BMNI-16 (Babesia microti)
CN Antigen BMNI-16 (Babesia microtia fragment)
CI MAN

```
SQL  128

SEQ      1 LWFIKMVSFK SILVPYITLF LMSGAVFASD TDPEAGGPSE AGGPSGTVGP
                                             ======= ==========
        51 SEAGGPSEAG GPSGTGWPSE AGGPSEAGGP SEAGGPSEAG GPSGTGWPSG
           ========== ========== ========== ========== ==========
       101 TGWPSEAGWS SERFGYQLLP YSRRIVIF
           ========== =
HITS AT:   34-111

REFERENCE   1:  135:367756

REFERENCE   2:  134:294507

REFERENCE   3:  134:158493

REFERENCE   4:  133:295360

REFERENCE   5:  128:292989

L5   ANSWER 25 OF 32  REGISTRY  COPYRIGHT 2002 ACS
RN   206205-21-2  REGISTRY
CN   Antigen (Babesia microtia 267-amino acid) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   247: PN: WO0185947 SEQID: 28 claimed protein
CN   28: PN: US6183976 SEQID: 28 claimed protein
CN   28: PN: US6214971 SEQID: 28 claimed protein
CN   28: PN: WO0060090 SEQID: 28 unclaimed protein
CN   Antigen BMNI-13 (Babesia microti)
CN   Antigen BMNI-13 (Babesia microtia fragment)
CN   Antigen BMNI-13 (Bombesia microtia antigen BMNI-17 fragment)
CI   MAN
SQL  267

SEQ      1 LWFIKMVSFK SILVPYITLF LMSGAVFASD TDPEAGGPSG TVGPSEAGGP
                                             ======= ==========
        51 SEAGGPSGTG WPSEAGGPSE AGGPSGTGWP SEAGWSSERF GYQLLPYSRR
           ========== ========== ========== =======
       101 IVTFNEVCLS YIYKHSVMIL ERDRVNDGHK DYIEEKTKEK NKLKKELEKC
       151 FPEQYSLMKK EELARIFDNA STISSKYKLL VDEISNKAYG TLEGPAADNF
       201 DHFRNIWKSI VLKDMFIYCD LLLQHLIYKF YYDNTINDIK KNFDESKSKA
       251 LVLRDKITKK DVYVNDH
HITS AT:   34-87

REFERENCE   1:  135:367756

REFERENCE   2:  134:294507

REFERENCE   3:  134:158493

REFERENCE   4:  133:295360

REFERENCE   5:  131:85409

REFERENCE   6:  128:292989

L5   ANSWER 26 OF 32  REGISTRY  COPYRIGHT 2002 ACS
RN   206205-20-1  REGISTRY
```

CN Antigen (Babesia microtia 121-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 246: PN: WO0185947 SEQID: 27 claimed protein
CN 27: PN: US6183976 SEQID: 27 claimed protein
CN 27: PN: US6214971 SEQID: 27 claimed protein
CN 27: PN: WO0060090 SEQID: 27 unclaimed protein
CN Antigen BMNI-12 (Babesia microti)
CN Antigen BMNI-12 (Babesia microtia fragment)
CN Antigen BMNI-12 (Bombesia microtia antigen BMNI-17 fragment)
CI MAN
SQL 121

```
SEQ      1 LWFIKMVSFK SILVPYITLF LMSGAVFASD TDPEAGGPSE AGGPSEAGGP
                                              ======= ==========
        51 SGTVGPSEAG GPSEAGGPSG TGWPSEAGGP SEAGGPSGTG WPSEAGWSSE
           ========== ========== ========== ========== =========
       101 RFGYQLLPYS RRIVTFNEVC L
```

HITS AT: 34-99

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 27 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-16-5** REGISTRY
CN Antigen (Babesia microtia 303-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 23: PN: US6183976 SEQID: 23 claimed protein
CN 23: PN: US6214971 SEQID: 23 claimed protein
CN 23: PN: WO0060090 SEQID: 23 unclaimed protein
CN 242: PN: WO0185947 SEQID: 23 claimed protein
CN Antigen BMNI-6 (Babesia microti)
CN Antigen BMNI-6 (Babesia microtia fragment)
CN Antigen BMNI-6 (Bombesia microtia antigen BMNI-17 fragment)
CI MAN
SQL 303

```
SEQ      1 LWFIKMVSFK SILVPYITLF LMSGAVFAGD TDREAGGPSG TVGPSEAGGP
                                              ======= ==========
        51 SEAGGPSEAG GPSEAGGPSE AGGPSEAGGP SEAGGPSEAG GPSGTGWPSE
           ========== ========== ========== ========== ==========
       101 AGWPSEAGWP SEAGWPSEAG WPSEAGWPSE RFGYQLLWYS RRIVIFNEIY
           ========== ========== ==========
       151 LSHIYEHSVM ILERDRVNDG HKDYIEEKTK EKNKLKKELE KCFPEQYSLM
       201 KKEELARIID NASTISSKYK LLVDEISNKA YGTLEGPAAD DFDHFRNIWK
       251 SIVPKNMFLY CDLLLKHLIR KFYCDNTIND IKKNFDDIEK LGCFQARSFL
       301 PVN
```

HITS AT: 34-129

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 28 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-13-2** REGISTRY
CN Antigen (Babesia microtia 367-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 20: PN: US6183976 SEQID: 20 claimed protein
CN 20: PN: US6214971 SEQID: 20 claimed protein
CN 20: PN: WO0060090 SEQID: 20 unclaimed protein
CN 239: PN: WO0185947 SEQID: 20 claimed protein
CN Antigen BMNI-3 (Babesia microti)
CN Antigen BMNI-3 (Babesia microtia fragment)
CN Antigen BMNI-3 (Bombesia microtia antigen BMNI-17 fragment)
CI MAN
SQL 367

SEQ 1 LWFIKMVSFK SILVPYITLF LMSGAVFASD TDPEAGGPSE AGGPSGTVGP
=======
51 SEAGGPSEAG GPSGTGWPSE AGGPSEAGGP SEAGGPSEAG GPSGTGWPSG
101 TGWPSEAGWS SERFGYQLLP YSRRIVIFNE VCLSYIYKHS VMILERDRVN
=
151 DGHKDYIEEK TKEKNKLKKE LEKCFPEQYS LMKKEELARI FDNASTISSK
201 YKLLVDEISN KAYGTLEGPA ADNFDHFRNI WKSIVLKDMF IYCDLLLQHL
251 IYKFYYDNTV NDIKKNFDES KSKALVLRDK ITKKDGDYNT HFEDMIKELN
301 SAAEEFNKIV DIMISNIGDY DEYDSIASFK PFLSMITEIT KITKVSNVII
351 PGIKALTLTV FLIFITK
HITS AT: 34-111

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 29 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-12-1** REGISTRY
CN Antigen (Babesia microtia 310-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 19: PN: US6183976 SEQID: 19 claimed protein
CN 19: PN: US6214971 SEQID: 19 claimed protein
CN 19: PN: WO0060090 SEQID: 19 unclaimed protein
CN 238: PN: WO0185947 SEQID: 19 claimed protein

CN Antigen BMNI-2 (Babesia microti)
CN Antigen BMNI-2 (Babesia microtia fragment)
CN Antigen BMNI-2 (Bombesia microtia antigen BMNI-17 fragment)
CI MAN
SQL 310

SEQ 1 MSGAVFASDT DPEAGGPSEA GGPSGTVGPS EAGGPSEAGG PSGTVGPSEA
51 GGPSEAGGPS GTGWPSEAGG PSEAGGPSGT VGPSEAGGPS EAGGPSGTGW
101 PSEAGGPSEA GGPSEAGGPS EAGGPSGTGW PSGTGWPSEA GWSSERFGYQ
151 LLPYSRRIVI FNEVCLSYIY KHSVMILERD RVNDGHKDYI EEKTKEKNKL
201 KKELEKCFPE QYSLMKKEEL ARIFDNASTI SSKYKLLVDE ISNKAYGTLE
251 GPAADNFDHF RNIWKSIVLK DMFIYCDLLL QHLIYKFYYD NTVNDIKKNF
301 DESWTQTLKE
HITS AT: 13-144

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 30 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **206205-11-0** REGISTRY
CN Antigen (Babesia microtia 263-amino acid) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 18: PN: US6183976 SEQID: 18 claimed protein
CN 18: PN: US6214971 SEQID: 18 claimed protein
CN 18: PN: WO0060090 SEQID: 18 unclaimed protein
CN 237: PN: WO0185947 SEQID: 18 claimed protein
CN Antigen BMNI-1 (Babesia microti)
CN Antigen BMNI-1 (Babesia microtia fragment)
CN Antigen BMNI-1 (Bombesia microtia antigen BMNI-17 fragment)
CI MAN
SQL 263

SEQ 1 LFLMSGAVFA SDTDPEAGGP SEAGGPSGTV GPSEAGGPSE AGGPSGTGWP
51 SEAGGPSEAG GPSEAGGPSE AGGPSGTGWP SGTGWPSEAG WSSERFGYQL
101 LPYSRRIVIF NEVCLSYIYK HSVMILERDR VNDGHKDYIE EKTKEKNKLK
151 KELEKCFPEQ YSLMKKEELA RIFDNASTIS SKYKLLVDEI SNKAYGTLEG
201 PAADNFDHFR NIWKSIVLKD MFIYCDLLLQ HLIYKFYYDN TVNDIKKNFD
251 ESKSKALVLR DKI
HITS AT: 16-93

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 31 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **205488-54-6** REGISTRY
CN L-Serine, L-.alpha.-glutamyl-L-alanylglycylglycyl-L-prolyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-prolyl-L-serylglycyl-L-threonylglycyl-L-tryptophyl-L-prolyl-L-seryl-L-.alpha.-glutamyl-L-alanylglycyl-L-tryptophylglycyl-L-seryl-L-.alpha.-glutamyl-L-alanylglycyl-L-tryptophyl-L-seryl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 17: PN: WO0185947 SEQID: 48 unclaimed sequence
CN 48: PN: US6183976 SEQID: 48 claimed sequence
CN 48: PN: US6214971 SEQID: 48 unclaimed sequence
CN 89: PN: WO0060090 SEQID: 48 unclaimed sequence
CN Antigen BABS-2 (Babesia microtia fragment)
SQL 30

SEQ 1 EAGGPSGTVG PSGTGWPSEA GWGSEAGWSS
========== ======== ======
HITS AT: 1-18, 25-30

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

L5 ANSWER 32 OF 32 REGISTRY COPYRIGHT 2002 ACS
RN **205488-48-8** REGISTRY
CN L-Serine, L-seryl-L-.alpha.-glutamyl-L-alanylglycylglycyl-L-prolyl-L-seryl-L-.alpha.-glutamyl-L-alanylglycylglycyl-L-prolyl-L-serylglycyl-L-threonylglycyl-L-tryptophyl-L-threonyl-L-serylglycyl-L-threonylglycyl-L-tryptophyl-L-prolyl-L-seryl-L-.alpha.-glutamyl-L-alanylglycyl-L-tryptophyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 16: PN: WO0185947 SEQID: 47 unclaimed sequence
CN 47: PN: US6183976 SEQID: 47 claimed sequence
CN 47: PN: US6214971 SEQID: 47 unclaimed sequence
CN 88: PN: WO0060090 SEQID: 47 unclaimed sequence
CN Antigen BABS-1 (Babesia microtia fragment)
SQL 30

SEQ 1 SEAGGPSEAG GPSGTGWTSG TGWPSEAGWS
========= === = =====
HITS AT: 2-13, 20-25

REFERENCE 1: 135:367756

REFERENCE 2: 134:294507

REFERENCE 3: 134:158493

REFERENCE 4: 133:295360

REFERENCE 5: 131:85409

REFERENCE 6: 128:292989

FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 12:09:33 ON 30 AUG 2002)

L6 4373 S "REED S"?/AU — Author(s)

L8 215 SEA ABB=ON PLU=ON "LODES M"?/AU
L9 927 SEA ABB=ON PLU=ON "HOUGHTON R"?/AU
L10 177 SEA ABB=ON PLU=ON "SLEATH P"?/AU
L11 295 SEA ABB=ON PLU=ON "MCNEILL P"?/AU
L12 416 SEA ABB=ON PLU=ON "HOMER M"?/AU
L13 113 SEA ABB=ON PLU=ON "SECRIST H"?/AU
L14 2 SEA ABB=ON PLU=ON L6 AND L8 AND L9 AND L10 AND L11 AND L12 AND L13
L15 231 SEA ABB=ON PLU=ON L6 AND (L8 OR L9 OR L10 OR L11 OR L12 OR L13)
L16 115 SEA ABB=ON PLU=ON L8 AND (L9 OR L10 OR L11 OR L12 OR L13)
L17 85 SEA ABB=ON PLU=ON L9 AND (L10 OR L11 OR L12 OR L13)
L18 19 SEA ABB=ON PLU=ON L10 AND (L11 OR L12 OR L13)
L19 2 SEA ABB=ON PLU=ON L11 AND (L12 OR L13)
L20 2 SEA ABB=ON PLU=ON L12 AND L13
L21 43 SEA ABB=ON PLU=ON (L15 OR L16 OR L17 OR L6 OR L8 OR L9 OR L10 OR L11 OR L12 OR L13) AND MICROTI
L22 56 SEA ABB=ON PLU=ON L14 OR L18 OR L19 OR L20 OR L21
L23 27 DUP REM L22 (29 DUPLICATES REMOVED)

L23 ANSWER 1 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 1
ACCESSION NUMBER: 2002:275811 HCAPLUS
DOCUMENT NUMBER: 136:308523
TITLE: Compositions and methods for WT1 specific immunotherapy
INVENTOR(S): Gaiger, Alexander; **McNeill, Patricia D.**; Smithgall, Molly; Moulton, Gus; Vedvick, Thomas S.; **Sleath, Paul R.**; Mossman, Sally; Evans, Lawrence; Spies, A. Gregory; Boydston, Jeremy
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: PCT Int. Appl., 260 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002028414	A1	20020411	WO 2001-US31139	20011003

WO 2002028414 B1 20020718
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG
PRIORITY APPLN. INFO.: US 2000-684361 A 20001006
US 2000-685830 A 20001009
US 2001-785019 A 20010215
US 2001-938864 A 20010824

AB Compns. and methods for the therapy of malignant diseases, such as leukemia and cancer, are disclosed. The compns. comprise one or more of a WT1 polynucleotide, a WT1 polypeptide, an antigen-presenting cell presenting a WT1 polypeptide, an antibody that specifically binds to a WT1 polypeptide; or a T cell that specifically reacts with a WT1 polypeptide. Such compns. may be used, for example, for the prevention and treatment of metastatic diseases.

REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 2 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 2
ACCESSION NUMBER: 2001:833517 HCAPLUS
DOCUMENT NUMBER: 135:367756
TITLE: Babesia **microti** antigens and methods for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): **Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond L.; Sleath, Paul R.; McNeill, Patricia D.; Homer, Mary J.; Secrist, Heather**
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: PCT Int. Appl., 195 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001085947	A2	20011115	WO 2001-US15192	20010509

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH,

CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

US 2001029295	A1	20011011	US 2000-737178		20001213
PRIORITY APPLN. INFO.:			US 2000-569098	A	20000510
			US 2000-605724	A	20000627
			US 2000-656688	A	20000907
			US 2000-685436	A	20001010
			US 2000-737178	A	20001213
			US 2001-794764	A	20010226
			US 1996-723142	A2	19961001
			US 1997-845258	A2	19970424
			US 1997-990571	A2	19971211
			WO 1998-US26437	A2	19981211
			US 1999-286488	A2	19990405
			US 2000-528784	A2	20000317
			WO 2000-US9136	A2	20000405

AB Compds. and methods for the diagnosis and treatment of B. **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and immunogenic compns. comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

L23 ANSWER 3 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 3

ACCESSION NUMBER:	2001:748300 HCAPLUS
DOCUMENT NUMBER:	135:299589
TITLE:	Nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S):	**Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond L.; Sleath, Paul R.; McNeill, Patricia D.**
PATENT ASSIGNEE(S):	USA
SOURCE:	U.S. Pat. Appl. Publ., 26 pp., Cont.-in-part of U. S. Ser. No. 685,436. CODEN: USXXCO
DOCUMENT TYPE:	Patent
LANGUAGE:	English
FAMILY ACC. NUM. COUNT:	7

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2001029295	A1	20011011	US 2000-737178	20001213
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424
US 6214971	B1	20010410	US 1997-990571	19971211
WO 9929869	A1	19990617	WO 1998-US26437	19981211

W: AU, CA, JP, MX, NZ
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE

WO 2000060090 A1 20001012 WO 2000-US9136 20000405
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
WO 2001085947 A2 20011115 WO 2001-US15192 20010509
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.:

US	1996-723142	A2 19961001
US	1997-845258	A2 19970424
US	1997-990571	A2 19971211
WO	1998-US26437	A2 19981211
US	1999-286488	A2 19990405
US	2000-528784	A2 20000317
WO	2000-US9136	A2 20000405
US	2000-569098	A2 20000510
US	2000-605724	A2 20000627
US	2000-656688	A2 20000907
US	2000-685436	A2 20001010
US	2000-737178	A 20001213
US	2001-794764	A 20010226

AB Compds. and methods for the diagnosis and treatment of B. microtiinfection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. microtiantigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and immunogenic compns. comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. microtiinfection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

L23 ANSWER 4 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 4
ACCESSION NUMBER: 2001:255942 HCAPLUS
DOCUMENT NUMBER: 134:294507
TITLE: Compounds and methods for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): **Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond**
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: U.S., 78 pp., Cont.-in-part of U.S. Ser. No. 845,258.
CODEN: USXXAM

DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6214971	B1	20010410	US 1997-990571	19971211
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424
WO 9929869	A1	19990617	WO 1998-US26437	19981211
W: AU, CA, JP, MX, NZ				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
AU 9918204	A1	19990628	AU 1999-18204	19981211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 1996-723142	A2	19961001
US 1997-845258	A2	19970424
US 1997-990571	A	19971211
WO 1998-US26437	W	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405
US 2000-569098	A2	20000510
US 2000-605724	A2	20000627
US 2000-656688	A2	20000907
US 2000-685436	A2	20001010

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

REFERENCE COUNT: 13 THERE ARE 13 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 5 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 5
ACCESSION NUMBER: 2001:91448 HCAPLUS
DOCUMENT NUMBER: 134:158493
TITLE: Nucleic acids and proteins for the diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): **Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond; Sleath, Paul R.**
PATENT ASSIGNEE(S): Corixa Corporation, USA
SOURCE: U.S., 66 pp., Cont.-in-part of U.S. Ser. No. 723,142.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6183976	B1	20010206	US 1997-845258	19970424
US 6306396	B1	20011023	US 1996-723142	19961001
EP 834567	A2	19980408	EP 1997-117067	19971001
EP 834567	A3	19990331		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6214971	B1	20010410	US 1997-990571	19971211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 1996-723142	A2	19961001
US 1997-845258	A	19970424
US 1997-990571	A2	19971211
WO 1998-US26437	A2	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405
US 2000-569098	A2	20000510
US 2000-605724	A2	20000627
US 2000-656688	A2	20000907
US 2000-685436	A2	20001010

OTHER SOURCE(S): MARPAT 134:158493

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. The DNA sequences encoding B. **microti** antigens were prepd. by screening a B. **microti** expression library with sera obtained from patients infected with B. **microti**. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

REFERENCE COUNT: 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 6 OF 27 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:6260 BIOSIS
DOCUMENT NUMBER: PREV200200006260
TITLE: Compounds and methods for the diagnosis and treatment of B. **microti** infection.
AUTHOR(S): **Reed, Steven G.; Lodes, Michael J. ; Houghton, Raymond; Sleath, Paul R.**
ASSIGNEE: Corixa Corporation
PATENT INFORMATION: US 6306396 October 23, 2001
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Oct. 23, 2001) Vol. 1251, No. 4, pp. No Pagination. e-file.
ISSN: 0098-1133.
DOCUMENT TYPE: Patent

LANGUAGE: English

AB Compounds and methods for the diagnosis and treatment of B. **microti** infection are disclosed. The compounds provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compositions and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits containing such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biological samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

L23 ANSWER 7 OF 27 WPIDS (C) 2002 THOMSON DERWENT

ACCESSION NUMBER: 2002-062250 [08] WPIDS

CROSS REFERENCE: 1998-609891 [51]; 2000-160675 [11]

DOC. NO. NON-CPI: N2002-046184

DOC. NO. CPI: C2002-017807

TITLE: Novel polynucleotide encoding polypeptides useful for detecting Ehrlichia infection in patients and biological samples, and for treating human granulocytic ehrlichiosis, comprise an Ehrlichia antigen.

DERWENT CLASS: B04 C06 D16 S03

INVENTOR(S): **HOUGHTON, R L; LODES, M J; MCNEILL, P D; REED, S G**

PATENT ASSIGNEE(S): (CORI-N) CORIXA CORP; (HOUG-I) HOUGHTON R L; (LODE-I) LODES M J; (MCNE-I) MCNEILL P D; (REED-I) REED S G

COUNTRY COUNT: 95

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2001085949	A2	20011115	(200208)*	EN	132
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW					
AU 2001059507	A	20011120	(200219)		
US 2002068343	A1	20020606	(200241)		

APPLICATION DETAILS:

PATENT NO	KIND		APPLICATION	DATE
WO 2001085949	A2		WO 2001-US14518	20010504
AU 2001059507	A		AU 2001-59507	20010504
US 2002068343	A1	CIP of	US 1997-821324	19970321
		CIP of	US 1997-975762	19971120
		CIP of	US 1998-106582	19980629
		CIP of	US 1998-159469	19980923
		CIP of	US 1999-295028	19990420
		CIP of	US 2000-566617	20000508

CIP of	US 2000-693542	20001020
	US 2001-798042	20010302

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2001059507 A	Based on	WO 200185949
US 2002068343 A1	CIP of	US 6207169
	CIP of	US 6231869
	CIP of	US 6277381
	CIP of	US 6306402

PRIORITY APPLN. INFO: US 2001-798042 20010302; US 2000-566617 20000508; US 2000-693542 20001020; US 1997-821324 19970321; US 1997-975762 19971120; US 1998-106582 19980629; US 1998-159469 19980923; US 1999-295028 19990420

AN 2002-062250 [08] WPIDS

CR 1998-609891 [51]; 2000-160675 [11]

AB WO 200185949 A UPAB: 20020701

NOVELTY - An isolated polynucleotide (I) comprising a sequence (S1) chosen from 36 nucleotides of defined bp fully in the specification such as 1345, 1132, 554, 559, 201, 467, 530, 1185, 1131, 800, 1011, 513, 464, 527, 464, 860, 484 or 1039 bp, complement of (S1), a sequence hybridizable under moderate stringent conditions to (S1), a sequence which is 75% or 90% identical to (S1) or degenerate variant of (S1), is new.

DETAILED DESCRIPTION - (I) is chosen from the determined DNA sequence of human granulocytic ehrlichiosis (HGE)-1, HGE-3, HGE-6, HGE-8, HGE-11-13, HGE-23, HGE-24, the 5' DNA sequence of HGE-7, HGE-2, HGE-9, HGE-14-18, HGE-25, extended DNA sequences of HGE-2, HGE-7, HGE-8, HGE-11, HGE-14-16, HGE-18, HGE-23, HGE-25, the determined 3' DNA sequence of HGE-17, the full-length cDNA sequence for HGE-17, corrected cDNA sequence for HGE-14, HGE-1, and the reverse complement of HGE-2, HGE-14 and HGE-15. INDEPENDENT CLAIMS are also included for the following:

(1) an isolated polypeptide (II) comprising a sequence encoded by (I) or a sequence 70% or 90% identical to a sequence encoded by (I);

(2) an isolated antigenic epitope (III) of an Ehrlichia antigen comprising a sequence of 41 or 125 amino acids defined in the specification;

(3) an isolated polypeptide comprising at least two antigenic epitopes as above;

(4) a recombinant expression vector (IV) comprising (I);

(5) a host cell (V) transformed with (IV);

(6) a fusion protein (VI) comprising (II) or (III);

(7) a diagnostic kit comprising (II), (III) and (VI) and a detection reagent;

(8) a diagnostic kit comprising at least two oligonucleotide probes or primer specific for (I);

(9) an isolated antibody (VII), or its antigen binding fragment that specifically binds to (II) or (III);

(10) a composition (VIII) comprising any one chosen from (I)-(III), (VI) and (VII), and immunostimulants;

(11) detecting Ehrlichia infection in a biological sample, by contacting the sample with oligonucleotide primers or probes, or a

binding agent capable of binding to (II), (III) or (VI) and detecting in the sample the presence of amplified polynucleotide sequence, polynucleotide sequence that hybridizes to the probe or polypeptide that binds to the binding agent; and

(12) detecting a disorder of Ehrlichia infection, Lyme disease and B.**microti** infection in a patient, by contacting the biological sample with (II), a Lyme disease antigen, and a B. **microti** antigen, and detecting the presence of antibodies in the biological sample that bind to either the polypeptide, Lyme disease antigen or the B.**microti** antigen.

ACTIVITY - Antibiotic. No supporting data is provided.

MECHANISM OF ACTION - Vaccine. No supporting data is given.

USE - (II), (III) and (VI) are useful for detecting Ehrlichia infection in a patient. (VIII) is useful for stimulating an immune response in a patient, for treating Ehrlichia infection in a patient (claimed). (II) is useful for serodiagnosis and treatment of human granulocytic ehrlichiosis (HGE). (VII) is useful in diagnostic test to detect the presence of Ehrlichia antigens and for detecting Ehrlichia infection in a patient.

Dwg.0/2

L23 ANSWER 8 OF 27 SCISEARCH COPYRIGHT 2002 ISI (R)
ACCESSION NUMBER: 2001:943909 SCISEARCH
THE GENUINE ARTICLE: 494QP
TITLE: Innate resistance to Babesia infection is influenced by genetic background and gender
AUTHOR: Aguilar-Delfin I; **Homer M J**; Wettstein P J; Persing D H (Reprint)
CORPORATE SOURCE: Corixa Corp, Suite 200, 1124 Columbia St, Seattle, WA 98104 USA (Reprint); Mayo Clin & Mayo Fdn, Dept Immunol, Rochester, MN 55905 USA; Corixa Corp, Seattle, WA 98104 USA; Infect Dis Res Inst, Seattle, WA 98104 USA
COUNTRY OF AUTHOR: USA
SOURCE: INFECTION AND IMMUNITY, (DEC 2001) Vol. 69, No. 12, pp. 7955-7958.
Publisher: AMER SOC MICROBIOLOGY, 1752 N ST NW, WASHINGTON, DC 20036-2904 USA.
ISSN: 0019-9567.
DOCUMENT TYPE: Article; Journal
LANGUAGE: English
REFERENCE COUNT: 30

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Infection of severe combined immunodeficient mice with Babesia sp. strain WA1 was studied to assess the contributions of innate and adaptive immunity in resistance to acute babesiosis. The scid mutation showed little effect in genetically susceptible C3H mice and did not decrease the inherent resistance of C57BL/6 mice to the infection, suggesting that innate immunity plays a central role in determining the course of Babesia infection in these strains. In contrast, the scid mutation dramatically impaired resistance in moderately susceptible BALB/c mice, suggesting that acquired immunity may play an important secondary role. In comparison to their female counterparts, male mice of different genetic backgrounds showed increased resistance to the infection, indicating that the gender of the host may influence protection against babesiosis.

L23 ANSWER 9 OF 27 SCISEARCH COPYRIGHT 2002 ISI (R)
ACCESSION NUMBER: 2001:546315 SCISEARCH
THE GENUINE ARTICLE: 447RN
TITLE: Serodiagnosis of human granulocytic ehrlichiosis by using novel combinations of immunoreactive recombinant proteins
AUTHOR: **Lodes M J (Reprint)**; Mohamath R; Reynolds L D; **McNeill P**; Kolbert C P; Bruinsma E S; Benson D R; Hofmeister E; **Reed S G**; **Houghton R L**; Persing D H
CORPORATE SOURCE: Corixa Corp, 1124 Columbia St, Suite 200, Seattle, WA 98104 USA (Reprint); Infect Dis Res Inst, Seattle, WA 98104 USA; Corixa Corp, Seattle, WA 98104 USA; Univ Washington, Dept Pathobiol, Seattle, WA 98195 USA; Mayo Clin & Mayo Fdn, Dept Lab Med & Pathol, Rochester, MN 55905 USA
COUNTRY OF AUTHOR: USA
SOURCE: JOURNAL OF CLINICAL MICROBIOLOGY, (JUL 2001) Vol. 39, No. 7, pp. 2466-2476.
Publisher: AMER SOC MICROBIOLOGY, 1752 N ST NW, WASHINGTON, DC 20036-2904 USA.
ISSN: 0095-1137.
DOCUMENT TYPE: Article; Journal
LANGUAGE: English
REFERENCE COUNT: 44

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB A panel of seven recombinant antigens, derived from Ehrlichia phagocytophila (the agent of human granulocytic ehrlichiosis), was evaluated by class-specific enzyme-linked immunosorbent assays (ELISAs) for utility in the diagnosis of the infection. Fourteen genomic fragments, obtained by serologic expression screening, contained open reading frames (ORFs) encoding 16 immunodominant antigens, Eleven of these antigens were members of the major surface protein (MSP) multigene family, Alignment of their predicted protein sequences revealed a pattern of conserved sequences, which contained short direct repeats, flanking a variable region. In addition, two genomic clones contained two and three MSP ORFs, respectively, indicating that these genes are clustered in tandem copies. The implications for this pattern of both genomic and protein arrangements in antigenic variations of MSPs and in their utilities in a diagnostic assay are discussed. In addition to two :MSP recombinant antigens (rHGE-1 and -3) and a fusion protein of these antigens (rErf-1), five further recombinants were evaluated by ELISA. Two of these antigens (rHGE-14 and -15) were novel, while a third (rHGE-2), with no known function, has been described. The final two recombinant antigens (rHGE-9 and -17) represent overlapping segments of the ankyrin gene (ank), The addition of rHGE-9 ELISA data resulted in the detection of 78% (21 of 27) of acute-phase sera. When serologic data for all recombinants are combined, 96.2% (26 of 27) of convalescent-phase patient serum samples and 85.2% (23 of 27) of acute-phase patient serum samples are detected, indicating the potential of these antigens for use in the development of a rapid serologic assay for the detection of E. phagocytophila infection.

L23 ANSWER 10 OF 27 SCISEARCH COPYRIGHT 2002 ISI (R)
ACCESSION NUMBER: 2001:818201 SCISEARCH
THE GENUINE ARTICLE: 472TV

TITLE: Peptide ELISA for detection of antibodies to Babesia **microti** in serum.
AUTHOR: **Houghton R L (Reprint)**; **Homer M J** ; Reynolds L D; **Sleath P C**; Cable R G; Militscher J E; **Lodes M J**; Berardi V; Leiby D A; Persing D H
CORPORATE SOURCE: Corixa Corp, Seattle, WA USA; Amer Red Cross, Farmington, CT USA; Amer Red Cross, Rockville, MD USA; Imugen, Norwood, MA USA
COUNTRY OF AUTHOR: USA
SOURCE: TRANSFUSION, (SEP 2001) Vol. 41, No. 9, Supp. [S], pp. 13S-13S.
Publisher: AMER ASSOC BLOOD BANKS, 8101 GLENBROOK RD, BETHESDA, MD 20814-2749 USA.
ISSN: 0041-1132.
DOCUMENT TYPE: Conference; Journal
LANGUAGE: English
REFERENCE COUNT: 0

L23 ANSWER 11 OF 27 SCISEARCH COPYRIGHT 2002 ISI (R)
ACCESSION NUMBER: 2001:818200 SCISEARCH
THE GENUINE ARTICLE: 472TV
TITLE: Evidence for transmission of Babesia **microti** from Connecticut blood donors to recipients
AUTHOR: Cable R G (Reprint); Badon S; Trouem-Trend J; Militscher J E; **Houghton R L**; **Lodes M J**; Persing D H; Eberhard M L; Pleniazek N J; Herwaldt B L; Leiby D A
CORPORATE SOURCE: Amer Red Cross Blood Serv, Farmington, CT USA; Amer Red Cross, Holland Lab, Rockville, MD USA; Corixa Corp, Seattle, WA USA; Ctr Dis Control & Prevent, Atlanta, GA USA
COUNTRY OF AUTHOR: USA
SOURCE: TRANSFUSION, (SEP 2001) Vol. 41, No. 9, Supp. [S], pp. 12S-13S.
Publisher: AMER ASSOC BLOOD BANKS, 8101 GLENBROOK RD, BETHESDA, MD 20814-2749 USA.
ISSN: 0041-1132.
DOCUMENT TYPE: Conference; Journal
LANGUAGE: English
REFERENCE COUNT: 0

L23 ANSWER 12 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 6
ACCESSION NUMBER: 2000:725776 HCAPLUS
DOCUMENT NUMBER: 133:295360
TITLE: Antigens of Babesia microtia for use in the diagnosis, prophylaxis, and treatment of babesiosis
INVENTOR(S): Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond L.; **Sleath, Paul R.**; **Mcneill, Patricia D.**
PATENT ASSIGNEE(S): Corixa Corp., USA
SOURCE: PCT Int. Appl., 118 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000060090	A1	20001012	WO 2000-US9136	20000405
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
EP 1169455	A1	20020109	EP 2000-921771	20000405
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:	US 1999-286488	A	19990405
	US 2000-528784	A	20000317
	US 1996-723142	A2	19961001
	US 1997-845258	A2	19970424
	US 1997-990571	A2	19971211
	WO 1998-US26437	A2	19981211
	WO 2000-US9136	W	20000405
	US 2000-569098	A2	20000510
	US 2000-605724	A2	20000627
	US 2000-656688	A2	20000907
	US 2000-685436	A2	20001010

OTHER SOURCE(S): MARPAT 133:295360

AB Compds. and methods for the diagnosis and treatment of B. microtia infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. microtia antigen and DNA sequences encoding such polypeptides. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. microtia infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided. Cloning of antigen genes by antibody screening of expression libraries with antiserum. Seventeen antigens were identified and several of these showed common sequences. The clones also contained telomere repeat sequences indicating that they were located near the telomere. Use of the antigens in diagnostic detection of B. microtia is demonstrated.

REFERENCE COUNT: 2 THERE ARE 2 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 13 OF 27 WPIDS (C) 2002 THOMSON DERWENT
ACCESSION NUMBER: 2000-160675 [14] WPIDS
CROSS REFERENCE: 1998-609891 [51]; 2002-062250 [01]
DOC. NO. NON-CPI: N2000-119888
DOC. NO. CPI: C2000-050162
TITLE: New compounds and methods for the diagnosis of Ehrlichia infection, particularly Human granulocytic ehrilichiosis.

DERWENT CLASS: A96 B04 D16 S03
INVENTOR(S): **HOUGHTON, R L; LODES, M J; MCNEILL, P D; REED, S G**
PATENT ASSIGNEE(S): (CORI-N) CORIXA CORP; (HOUG-I) HOUGHTON R L; (LODE-I) LODES M J; (MCNE-I) MCNEILL P D; (REED-I) REED S G
COUNTRY COUNT: 85
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2000000615	A2	20000106	(200014)*	EN	108
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SL SZ UG ZW					
W: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZA ZW					
AU 9948474	A	20000117	(200026)		
US 6277381	B1	20010821	(200150)		
US 6306402	B1	20011023	(200165)		
EP 1144639	A2	20011017	(200169)	EN	
R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
US 2002064535	A1	20020530	(200240)		

APPLICATION DETAILS:

PATENT NO	KIND		APPLICATION	DATE
WO 2000000615	A2		WO 1999-US14793	19990629
AU 9948474	A		AU 1999-48474	19990629
US 6277381	B1	CIP of	US 1997-821324	19970321
		CIP of	US 1997-975762	19971120
		CIP of	US 1998-106582	19980629
		CIP of	US 1998-159469	19980923
			US 1999-295028	19990420
US 6306402	B1	CIP of	US 1997-821324	19970321
		CIP of	US 1997-975762	19971120
			US 1998-106582	19980629
EP 1144639	A2		EP 1999-932087	19990629
			WO 1999-US14793	19990629
US 2002064535	A1	CIP of	US 1997-821324	19970321
		CIP of	US 1997-975762	19971120
		Cont of	US 1998-106582	19980629
			US 1998-159469	19980923

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 9948474	A	Based on	WO 200000615
EP 1144639	A2	Based on	WO 200000615
US 2002064535	A1	CIP of	US 6207169
		CIP of	US 6231869
		Cont of	US 6306402

PRIORITY APPLN. INFO: US 1999-295028 19990420; US 1998-106582 19980629; US 1998-159469 19980923; US

Searcher : Shears 308-4994

1997-821324 19970321; US 1997-975762 19971120

AN 2000-160675 [14] WPIDS

CR 1998-609891 [51]; 2002-062250 [01]

AB WO 200000615 A UPAB: 20020701

NOVELTY - A polypeptide (P) comprising an immunogenic portion of an Ehrlichiaantigen or its variant that is encoded by one of 18 DNA sequences of 201-7091 base pairs (bp) (I)-(XVIII) (all sequences fully defined in the specification), their complements and DNA sequences that hybridize to sequences (I)-(XVIII) , are new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) An antigenic epitope (E) of an Ehrlichia antigen comprising an amino acid sequence of (XIX) or (XX) consisting of 41 and 125 amino acids (aa) respectively;

(2) A polypeptide (P') comprising at least two of (E);

(3) A DNA molecule (N) comprising a nucleotide sequence encoding (P) or (P');

(4) A recombinant expression vector (V) comprising (N);

(5) A host cell (H) transformed with (N);

(6) A fusion protein (F) comprising either at least one of (P) or (P') and/or at least one of (E) or a 376 aa sequence (XXI) and/or a 325 aa sequence (XXII).

(7) detecting Ehrlichia infection, Lyme disease and Babesia microtiinfection in a patient comprising:

(a) contacting a biological sample with at least one of (E), (P), (P') or (F) and a Lyme disease antigen and a B.**microti** antigen; and

(b) detecting the presence of antibodies that bind to (E), (P), (P') or (F) or the Lyme disease antigen or the B.**microti** antigen in the sample;

(8) A method similar to (7), comprising:

(a) contacting a biological sample with a specific binding agent to at least one of (E), (P), (P') or (F) or a Lyme disease antigen and a B.**microti** antigen; and

(b) detecting a polypeptide that binds to the binding agent, thereby detecting Ehrlichiainfection;

(9) A method similar to (7), comprising:

(a) contacting the sample with one or more probe oligonucleotides (or at least two primer oligonucleotides in a PCR reaction) where at least one is specific for (N); and

(b) detecting in the sample a DNA sequence that hybridizes to (or amplifies in the presence of) the oligonucleotide primers, thereby detecting Ehrlichia infection;

(10) A diagnostic kit (K) comprising:

(a) at least one of (P), (P'), (E) or (F); and

(b) a detection agent;

(11) A diagnostic kit (K') comprising at least two oligonucleotide primers or one oligonucleotide probe whereby at least one is specific for (N);

(12) A monoclonal antibody or polyclonal antibody that binds to (P), (P') or (E); and

(13) Vaccines comprising at least one of (P), (P'), (N) or (E) and a non-specific immune enhancer such as an adjuvant;

(14) A polypeptide comrpising an immunogenic portion of an Erlichia antigen associated with human granulocytic erlichiosis or its variant.

All sequences are fully defined in the specification.

USE - (P), (P'), (N), (F) and/or (E) are useful for the

detection and treatment of Ehrlichiainfection. (P), (P'), (F) and/or (E) can also be used to detect Lyme disease and B.microtiinfection. In particular, (P') can be used for the serodiagnosis and treatment of human granulocytic ehrlichiosis (HGE). Compositions of (P) or (P'), (N) and (E) can be used in the manufacture of a medicament for inducing protective immunity in a patient. The new vaccines are also used for inducing protective immunity in a patient.

ADVANTAGE - None given.

Dwg.0/2

L23 ANSWER 14 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 7
ACCESSION NUMBER: 2000:282570 HCAPLUS
DOCUMENT NUMBER: 133:72606
TITLE: Serological expression cloning of novel immunoreactive antigens of Babesia **microti**
AUTHOR(S): **Lodes, Michael J.; Houghton, Raymond L.**; Bruinsma, Elizabeth S.; Mohamath, Raodoh; Reynolds, Lisa D.; Benson, Darin R.; Krause, Peter J.; **Reed, Steven G.**; Persing, David H.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA
SOURCE: Infection and Immunity (2000), 68(5), 2783-2790
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Increased recognition of the prevalence of human babesiosis in the United States, together with rising concern about the potential for transmission of this infection by blood transfusion, has provided motivation to develop definitive serol. and mol. tests for the causative agent, Babesia **microti**. To develop more sensitive and specific assays for B. **microti**, the authors screened a genomic expression library with patient serum pools. This screening resulted in the identification of three classes of novel genes and an addnl. two novel, unrelated genes, which together encode a total of 17 unique B. **microti** antigens. The first class (BMN1-2 family) of genes encodes seven closely related antigens with a degenerate six-amino-acid repeat that shows limited homol. to Plasmodium sp. merozoite and sporozoite surface antigens. A second class (BMN1-8 family) of genes encodes six related antigens, and the third class (BMN1-17 family) of genes encodes two related antigens. The two remaining genes code for novel and unrelated sequences. Among the three classes of antigens and remaining novel sequences, five were chosen to code for the most immunodominant antigens (BMN1-2, -9, -15, and -17 and MN-10). Western blot anal. with the resulting recombinant proteins indicated that these antigens were targets of humoral immune responses during B. **microti** infection in humans.

REFERENCE COUNT: 36 THERE ARE 36 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 15 OF 27 MEDLINE DUPLICATE 8
ACCESSION NUMBER: 2000344709 MEDLINE
DOCUMENT NUMBER: 20344709 PubMed ID: 10885987
TITLE: Babesiosis.
AUTHOR: **Homer M J**; Aguilar-Delfin I; Telford S R

3rd; Krause P J; Persing D H
CORPORATE SOURCE: Corixa Corporation and The Infectious Disease Research Institute, Seattle, Washington 98104, USA.
SOURCE: CLINICAL MICROBIOLOGY REVIEWS, (2000 Jul) 13 (3) 451-69. Ref: 245
Journal code: 8807282. ISSN: 0893-8512.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200008
ENTRY DATE: Entered STN: 20000811
Last Updated on STN: 20000811
Entered Medline: 20000803

AB Babesiosis is an emerging, tick-transmitted, zoonotic disease caused by hematotropic parasites of the genus Babesia. Babesial parasites (and those of the closely related genus Theileria) are some of the most ubiquitous and widespread blood parasites in the world, second only to the trypanosomes, and consequently have considerable worldwide economic, medical, and veterinary impact. The parasites are intraerythrocytic and are commonly called piroplasms due to the pear-shaped forms found within infected red blood cells. The piroplasms are transmitted by ixodid ticks and are capable of infecting a wide variety of vertebrate hosts which are competent in maintaining the transmission cycle. Studies involving animal hosts other than humans have contributed significantly to our understanding of the disease process, including possible pathogenic mechanisms of the parasite and immunological responses of the host. To date, there are several species of Babesia that can infect humans, Babesia **microti** being the most prevalent. Infections with Babesia species generally follow regional distributions; cases in the United States are caused primarily by B. **microti**, whereas cases in Europe are usually caused by Babesia divergens. The spectrum of disease manifestation is broad, ranging from a silent infection to a fulminant, malaria-like disease, resulting in severe hemolysis and occasionally in death. Recent advances have resulted in the development of several diagnostic tests which have increased the level of sensitivity in detection, thereby facilitating diagnosis, expediting appropriate patient management, and resulting in a more accurate epidemiological description.

L23 ANSWER 16 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 9
ACCESSION NUMBER: 2000:74595 HCAPLUS
DOCUMENT NUMBER: 133:100272
TITLE: A polymorphic multigene family encoding an immunodominant protein from Babesia **microti**
AUTHOR(S): **Homer, M. J.**; Bruinsma, E. S.; **Lodes, M. J.**; Moro, M. H.; Telford, S., III; Krause, P. J.; Reynolds, L. D.; Mohamath, R.; Benson, D. R.; **Houghton, R. L.**; **Reed, S. G.**; Persing, D. H.
CORPORATE SOURCE: Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, 55905, USA
SOURCE: Journal of Clinical Microbiology (2000), 38(1),

362-368
CODEN: JCMIDW; ISSN: 0095-1137
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Human babesiosis in the United States is caused predominantly by Babesia **microti**, a tick-transmitted blood parasite. Improved testing methods for the detection of infection with this parasite are needed, since asymptomatic B. **microti** infection represents a potential threat to the blood supply in areas where B. **microti** is endemic. We performed immunoscreening of an expression library of genomic DNA from a human isolate of B. **microti** (strain MN1). Among 17 unique immunoreactive clones, we identified 9 which represent a related family of genes with little sequence homol. to other known sequences but with an architecture resembling that of several surface proteins of Plasmodium. Within this family, a tandem array of a degenerate six-amino-acid repeat (SEAGGP, SEAGWP, SGTGWP, SGTVGP) was found in various lengths between relatively well conserved segments at the N and C termini. In order to examine within-clone variation, we developed a PCR protocol for direct recovery of a specific bmn1-6 homolog directly from 30 human blood isolates, 4 corresponding hamster isolates, and 5 geog. corresponding Peromyscus leucopus (white-footed mouse) isolates. Isolates from the hamsters had the same sequences as those found in the corresponding human blood, suggesting that genetic variation of bmn1-6 does not occur during passage. However, clones from different patients were often substantially different from each other with regard to the no. and location of the degenerate repeats within the bmn1-6 homolog. Moreover, we found that strains that were closely related geog. were also closely related at the sequence level; nine patients, all from Nantucket Island, Mass., harbored clones that were indistinguishable from each other but that were distinct from those found in other northeastern or upper midwestern strains. We conclude that considerable genetic and antigenic diversity exists among isolates of B. **microti** from the United States and that geog. clustering of subtypes may exist. The nature of the bmn1-6 gene family suggests a mechanism of antigenic variation in B. **microti** that may occur by recombination, differential expression, or a combination of both mechanisms.

REFERENCE COUNT: 43 THERE ARE 43 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 17 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 10

ACCESSION NUMBER: 2000:87236 HCAPLUS
DOCUMENT NUMBER: 133:57223

TITLE: Multiepitope synthetic peptide and recombinant protein for the detection of antibodies to Trypanosoma cruzi in patients with treated or untreated Chagas' disease

AUTHOR(S): Houghton, Raymond L.; Benson, Darin R.; Reynolds, Lisa; **McNeill, Patricia**; **Sleath, Paul**; Lodes, Michael; Skeiky, Yasir A. W.; Badaro, Roberto; Krettli, Antoniana U.; Reed, Steven G.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA

SOURCE: Journal of Infectious Diseases (2000), 181(1),

325-330
CODEN: JIDIAQ; ISSN: 0022-1899
PUBLISHER: University of Chicago Press
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A tetrapeptide and a recombinant protein, each representing 4 immunodominant epitopes of Trypanosoma cruzi, were tested by use of ELISA for the detection of serum antibodies. Sera from individuals with Chagas' disease, including persons untreated and successfully or unsuccessfully treated, were tested. These assays detected antibody in 100% of the parasitemias. The antibody reactivity decreased based on the success of treatment. Higher sensitivity was obsd. for tetrapeptide/recombinant protein assays than for lysate-based ELISA, and specificity was improved, particularly with Leishmania sera. The results indicate that multiepitope antigens provide a more sensitive and specific alternative to lysate for detection of anti-T. cruzi antibodies, as required for developing blood screening assays.
REFERENCE COUNT: 29 THERE ARE 29 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 18 OF 27 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2000:394239 BIOSIS
DOCUMENT NUMBER: PREV200000394239
TITLE: Characterization of the Babesia **microti** chronic carrier state in a murine model.
AUTHOR(S): **Homer, M. J. (1)**; Bruinsma, E. S. (1); Aguilar-Delfin, I. (1); Moro, M. J. (1); Persing, D. H.
CORPORATE SOURCE: (1) Mayo Fndn, Rochester, MN USA
SOURCE: Abstracts of the General Meeting of the American Society for Microbiology, (2000) Vol. 100, pp. 283. print.
Meeting Info.: 100th General Meeting of the American Society for Microbiology Los Angeles, California, USA May 21-25, 2000 American Society for Microbiology . ISSN: 1060-2011.
DOCUMENT TYPE: Conference
LANGUAGE: English
SUMMARY LANGUAGE: English

L23 ANSWER 19 OF 27 SCISEARCH COPYRIGHT 2002 ISI (R)
ACCESSION NUMBER: 2000:888017 SCISEARCH
THE GENUINE ARTICLE: 364CN
TITLE: Serologic and nucleic acid evidence for babesia **microti** in Connecticut (CT) blood donors
AUTHOR: Leiby D A (Reprint); Chung A P; Triano L R; Cable R G; **Houghton R L; Lodes M J**
CORPORATE SOURCE: AMER RED CROSS, FARMINGTON, CT; CORIXA CORP, SEATTLE, WA; AMER RED CROSS, ROCKVILLE, MD
COUNTRY OF AUTHOR: USA
SOURCE: TRANSFUSION, (OCT 2000) Vol. 40, No. 10, Supp. [S], pp. S2-S2.
Publisher: AMER ASSOC BLOOD BANKS, 8101 GLENBROOK RD, BETHESDA, MD 20814-2749.
ISSN: 0041-1132.
DOCUMENT TYPE: Conference; Journal

FILE SEGMENT: LIFE; CLIN
LANGUAGE: English
REFERENCE COUNT: 0

L23 ANSWER 20 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 11
ACCESSION NUMBER: 1999:388301 HCAPLUS
DOCUMENT NUMBER: 131:85409
TITLE: Antigen and gene sequences for diagnosis and treatment of Babesia **microti** infection
INVENTOR(S): **Reed, Steven G.; Lodes, Michael J.; Houghton, Raymond; Sleath, Paul R.**; Persing, David; Bruinsma, Elizabeth
PATENT ASSIGNEE(S): Corixa Corporation, USA; Mayo Foundation for Medical Education and Research
SOURCE: PCT Int. Appl., 126 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9929869	A1	19990617	WO 1998-US26437	19981211
W: AU, CA, JP, MX, NZ				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 6214971	B1	20010410	US 1997-990571	19971211
AU 9918204	A1	19990628	AU 1999-18204	19981211
US 2001029295	A1	20011011	US 2000-737178	20001213

PRIORITY APPLN. INFO.:

US 1997-990571	A	19971211
US 1996-723142	A2	19961001
US 1997-845258	A2	19970424
WO 1998-US26437	W	19981211
US 1999-286488	A2	19990405
US 2000-528784	A2	20000317
WO 2000-US9136	A2	20000405
US 2000-569098	A2	20000510
US 2000-605724	A2	20000627
US 2000-656688	A2	20000907
US 2000-685436	A2	20001010

AB Compds. and methods for the diagnosis and treatment of Babesia **microti** infection are disclosed. The compds. provided include polypeptides that contain at least one antigenic portion of a B. **microti** antigen and DNA sequences encoding such polypeptides. Nine antigens share some homol., contain a degenerate repeat of six amino acids with 9-22 repeats occurring in each antigen, and bear some similarity to a Plasmodium falciparum merozoite surface antigen gene. A second group of five antigens bear some homol. to each other but do not show homol. to any previously identified sequences. Two synthetic peptides (BABS-1 and BABS-4) were made to the repeat region of isolated B. **microti** antigen BMNI-3. Twelve BMNI-6 homologs were obtained from hamster and human patients. Antigenic epitopes of such antigens are also provided, together with pharmaceutical compns. and vaccines comprising such polypeptides, DNA sequences or antigenic epitopes. Diagnostic kits contg. such polypeptides, DNA

sequences or antigenic epitopes and a suitable detection reagent may be used for the detection of B. **microti** infection in patients and biol. samples. Antibodies directed against such polypeptides and antigenic epitopes are also provided.

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 21 OF 27 MEDLINE DUPLICATE 12
ACCESSION NUMBER: 1999221786 MEDLINE
DOCUMENT NUMBER: 99221786 PubMed ID: 10203519
TITLE: Detection of enzootic babesiosis in baboons (Papio cynocephalus) and phylogenetic evidence supporting synonymy of the genera Entopolypoides and Babesia.
AUTHOR: Bronsdon M A; **Homer M J**; Magera J M; Harrison C; Andrews R G; Bielitzki J T; Emerson C L; Persing D H; Fritsche T R
CORPORATE SOURCE: Regional Primate Research Center, University of Washington School of Medicine, Seattle, Washington 98195, USA.
CONTRACT NUMBER: AI35191 (NIAID)
AI41103-01 (NIAID)
RR00166 (NCRR)
SOURCE: JOURNAL OF CLINICAL MICROBIOLOGY, (1999 May) 37 (5) 1548-53.
Journal code: 7505564. ISSN: 0095-1137.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AF081465
ENTRY MONTH: 199905
ENTRY DATE: Entered STN: 19990525
Last Updated on STN: 20000303
Entered Medline: 19990507

AB Blood smear evaluation of two baboons (Papio cynocephalus) experiencing acute hemolytic crises following experimental stem cell transplantation revealed numerous intraerythrocytic organisms typical of the genus Babesia. Both animals had received whole-blood transfusions from two baboon donors, one of which was subsequently found to display rare trophozoites of Entopolypoides macaci. An investigation was then undertaken to determine the prevalence of hematozoa in baboons held in our primate colony and to determine the relationship, if any, between the involved species. Analysis of thick and thin blood films from 65 healthy baboons (23 originating from our breeding facility, 26 originating from an out-of-state breeding facility, and 16 imported from Africa) for hematozoa revealed rare E. macaci parasites in 31%, with respective prevalences of 39, 35, and 12%. Phylogenetic analysis of nuclear small-subunit rRNA gene sequences amplified from peripheral blood of a baboon chronically infected with E. macaci demonstrated this parasite to be most closely related to Babesia **microti** (97.9% sequence similarity); sera from infected animals did not react in indirect fluorescent-antibody tests with Babesia **microti** antigen, however, suggesting that they represent different species. These results support an emerging view that the genus Entopolypoides Mayer 1933 is synonymous with that of the genus Babesia Starcovici 1893 and that the morphological variation noted

among intracellular forms is a function of alteration in host immune status. The presence of an underrecognized, but highly enzootic, Babesia sp. in baboons may result in substantial, unanticipated impact on research programs. The similarity of this parasite to the known human pathogen B. **microti** may also pose risks to humans undergoing xenotransplantation, mandating effective screening of donor animals.

L23 ANSWER 22 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 13
ACCESSION NUMBER: 1999:316369 HCAPLUS
DOCUMENT NUMBER: 131:143166
TITLE: A multi-epitope synthetic peptide and recombinant protein for the detection of antibodies to Trypanosoma cruzi in radioimmunoprecipitation-confirmed and consensus-positive sera
AUTHOR(S): Houghton, Raymond L.; Benson, Darin R.; Reynolds, Lisa D.; **McNeill, Patricia D.**; **Sleath, Paul R.**; Lodes, Michael J.; Skeiky, Yasir A. W.; Leiby, David A.; Badaro, Roberto; Reed, Steven G.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, 98104, USA
SOURCE: Journal of Infectious Diseases (1999), 179(5), 1226-1234
CODEN: JIDIAQ; ISSN: 0022-1899
PUBLISHER: University of Chicago Press
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Peptide epitopes of Trypanosoma cruzi have been identified through expression cloning. A tripeptide (2/D/E) contg. three epitopes (TcD, TcE, PEP-2) was used in ELISA to detect antibodies to T. cruzi in 239 of 240 consensus-pos. sera and 41 of 42 sera confirmed pos. by radioimmunopptn. assay. The 1 discrepant consensus-pos. serum was used to expression-clone a novel gene that contained a repeat sequence. A peptide corresponding to this sequence, TcLo1.2, was specific for T. cruzi. This antigen detected the discrepant consensus-pos. serum and enhanced reactivity of low-pos. sera in the tripeptide assay. A branched synthetic peptide, 2/D/E/Lo1.2, or a linear recombinant, r2/D/E/Lo1.2, realized all of the diagnostic features of the 4 epitopes, including the ability to boost reactivity of low-reactive sera. Thus, peptides and recombinants contg. multiple repeat epitopes are powerful tools for developing assays for T. cruzi antibody detection and have direct application in blood screening.
REFERENCE COUNT: 41 THERE ARE 41 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L23 ANSWER 23 OF 27 HCAPLUS COPYRIGHT 2002 ACS DUPLICATE 14
ACCESSION NUMBER: 1998:229027 HCAPLUS
DOCUMENT NUMBER: 128:292989
TITLE: Antigens of Babesia microtia and their use in the diagnosis and treatment of infection
INVENTOR(S): **Reed, Steven G.**; **Lodes, Michael J.**; **Houghton, Raymond**; **Sleath, Paul R.**
PATENT ASSIGNEE(S): Corixa Corp., USA
SOURCE: Eur. Pat. Appl., 113 pp.

CODEN: EPXXDW
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 7
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 834567	A2	19980408	EP 1997-117067	19971001
EP 834567	A3	19990331		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6306396	B1	20011023	US 1996-723142	19961001
US 6183976	B1	20010206	US 1997-845258	19970424
PRIORITY APPLN. INFO.:			US 1996-723142 A	19961001
			US 1997-845258 A	19970424

AB Antigens and epitopes of Babesia **microti** that can be used in the diagnosis and treatment of infection are described. Genes for the antigens or antibodies against them can be used in the detection of B. **microti**. CDNAS for these antigens were cloned by screening an expression library in .lambda.ZAP with antiserum to B. **microti**.

L23 ANSWER 24 OF 27 WPIDS (C) 2002 THOMSON DERWENT
ACCESSION NUMBER: 1998-609891 [51] WPIDS
CROSS REFERENCE: 2000-160675 [11]; 2002-062250 [01]
DOC. NO. CPI: C1998-182724
TITLE: Poly peptide(s) comprising immunogenic portion of Ehrlichia antigen - and encoding DNA sequences, useful for e.g. diagnosis and treatment of Ehrlichia infection, especially human granulocytic ehrlichiosis.
DERWENT CLASS: A96 B04 D16
INVENTOR(S): **HOUGHTON, R; LODES, M J; REED, S G; HOUGHTON, R L**
PATENT ASSIGNEE(S): (CORI-N) CORIXA CORP
COUNTRY COUNT: 81
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 9842740	A2	19981001	(199851)*	EN	140
RW: AT BE CH DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SZ UG ZW					
W: AL AM AT AU BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GE GH GM GW HU ID IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZW					
AU 9865794	A	19981020	(199909)		
EP 1007550	A2	20000614	(200033)	EN	
R: AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
US 6207169	B1	20010327	(200119)		
US 6231869	B1	20010515	(200129)		
JP 2002515763 W		20020528	(200238)		139

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE

WO 9842740	A2		WO 1998-US5695	19980323
AU 9865794	A		AU 1998-65794	19980323
EP 1007550	A2		EP 1998-911966	19980323
			WO 1998-US5695	19980323
US 6207169	B1	CIP of	US 1997-821324	19970321
			US 1997-975762	19971120
US 6231869	B1		US 1997-821324	19970321
JP 2002515763	W		JP 1998-545891	19980323
			WO 1998-US5695	19980323

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 9865794	A	Based on	WO 9842740
EP 1007550	A2	Based on	WO 9842740
JP 2002515763	W	Based on	WO 9842740

PRIORITY APPLN. INFO: US 1997-975762 19971120; US 1997-821324 19970321

AN 1998-609891 [51] WPIDS

CR 2000-160675 [11]; 2002-062250 [01]

AB WO 9842740 A UPAB: 20020701

A polypeptide comprising an immunogenic portion of an Ehrlichia antigen (or variant differing by conservative substitutions and/or modifications) is new, in which the antigen has amino acid sequence encoded by:

(i) one of nineteen 201-7091 bp sequences given in the specification, encoding Ehrlichia antigens;

(ii) sequences hybridising to (i); or

(iii) complements of (i) or (ii).

Also claimed are:

(1) an antigenic epitope of an Ehrlichia antigen with the 41 or 125 amino acid sequences (I) or (II) given in the specification, and a second polypeptide comprising at least two such antigenic epitopes.

(2) DNA molecules encoding polypeptides as above;

(3) expression vectors comprising (2);

(4) host cells (e.g. E. coli, yeast or mammalian) transformed with (3);

(5) fusion proteins comprising at least one of polypeptides and/or at least one of antigenic epitopes;

(6) diagnostic kits comprising at least one polypeptide/antigenic epitope/fusion protein (optionally immobilised on solid support e.g. nitrocellulose) plus detection reagent (optionally comprising reporter group e.g. radioisotope conjugated to binding agent e.g. anti-immunoglobulin; and

(7) monoclonal/polyclonal antibodies binding to polypeptides/antigenic epitope.

USE - The polypeptides are useful in the treatment of Ehrlichia infection, and as vaccines for the prevention of infection (claimed). They preferably comprise an immunogenic portion of an Ehrlichia antigen associated with human granulocytic ehrlichiosis (or a variant) (claimed) and are thus especially useful in the treatment of human granulocytic ehrlichiosis (HGE). The polypeptides, antigenic epitopes or DNA molecules can be combined with a suitable carrier in pharmaceutical compositions (claimed).

Such compositions and the vaccines of (7) are useful to manufacture medicaments for inducing protective immunity against Ehrlichia infection in patients (claimed), especially against HGE. The polypeptides/antigenic epitopes/fusion proteins are also useful to detect such infections in patients, by contacting biological samples (e.g. serum, and especially whole blood (claimed) with at least 1 polypeptide/antigenic epitope/fusion protein and detecting antibody binding (claimed). They can also be used to produce antibodies useful in diagnosis of such infections. The DNA molecules of (2) are similarly useful for diagnosing such infections (claimed). HGE is caused by a rodent bacterium normally transmitted to humans by the same tick which transmits Lyme disease and babesiosis. Co-infection with these diseases is thus possible, and the polypeptides/antigenic epitopes/fusion proteins may be used in methods to detect at least one of Ehrlichia infection, Lyme disease or B. **microti** infection in patients, by contacting samples (e.g. whole blood, saliva etc.) with at least one polypeptide/antigenic epitope/fusion protein, a Lyme disease antigen and a B. **microti** antigen, and detecting antibody binding (claimed); kits are provided (claimed).
Dwg.0/2

L23 ANSWER 25 OF 27 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 1998:37806 CONFSCI
DOCUMENT NUMBER: 98-037806
TITLE: Immunoreactivity of recombinant antigens of Babesia **microti** isolated using serological expression cloning
AUTHOR: **Houghton, R.L.**; Bruinsma, E.S.; Moro, M.H.; Krause, P.J.; Reynolds, L.D.; Mohamath, R.; Benson, D.R.; **Lodes, M.J.**; **Reed, S.G.**; Persing, D.H.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, USA
SOURCE: ASTMH, 60 Revere Drive, Suite 500, Northbrook, IL 60062, USA, Abstracts available. Price $10. Poster Paper No. 558.
Meeting Info.: 981 5000: 46th Annual Meeting of the American Society of Tropical Medicine and Hygiene (9815000). Lake Buena Vista, FL (USA). 7-11 Dec 1997. American Society of Tropical Medicine and Hygiene.
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: English

L23 ANSWER 26 OF 27 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 2002:14683 CONFSCI
DOCUMENT NUMBER: 02-014683
TITLE: Identification and partial characterization of secreted antigens from Babesia **microti** using a novel approach
AUTHOR: **Homer, M.J.**; **Lodes, M.J.**; Reynolds, L.D.; **Houghton, R.L.**; Persing, D.H.
CORPORATE SOURCE: Corixa Corporation, Seattle, WA, USA
SOURCE: American Society for Tropical Medicine, 60 Revere Dr., Suite 500, Northbrook, IL 60062, USA; phone: 847-480-9592; fax: 847-480-9282; email: astmh@astmh.org; URL: www.astmh.org. Poster Paper No.

671.
Meeting Info.: 000 5775: 50th Annual Meeting of the American Society for Tropical Medicine (0005775). Atlanta, GA (USA). 11-15 Nov 2001. Bill and Melinda Gates Foundation, Glaxo SmithKline, Oravax Inc., Berna Products.
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: English

L23 ANSWER 27 OF 27 CONFSCI COPYRIGHT 2002 CSA
ACCESSION NUMBER: 1998:37803 CONFSCI
DOCUMENT NUMBER: 98-037803
TITLE: Geographic variation within a gene encoding an immunoreactive protein from Babesia **microti**
AUTHOR: Bruinsma, E.S.; **Lodes, M.J.**; Moro, M.; Krause, P.J.; Reynolds, L.D.; Mohamath, R.; Benson, D.R.; **Houghton, R.L.**; **Reed, S.G.**; Persing, D.H.
CORPORATE SOURCE: Dep. Med. and Pathol., Mayo Clinic, Rochester, MN, USA
SOURCE: ASTMH, 60 Revere Drive, Suite 500, Northbrook, IL 60062, USA, Abstracts available. Price $10. Poster Paper No. 554.
Meeting Info.: 981 5000: 46th Annual Meeting of the American Society of Tropical Medicine and Hygiene (9815000). Lake Buena Vista, FL (USA). 7-11 Dec 1997. American Society of Tropical Medicine and Hygiene.
DOCUMENT TYPE: Conference
FILE SEGMENT: DCCP
LANGUAGE: English

FILE 'HOME' ENTERED AT 12:24:46 ON 30 AUG 2002

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 30, 2002, 13:18:44 ; Search time 51.8 Seconds
(without alignments)
12.866 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

X_1 = Glu
X_2 = Ala
X_3 = Gly

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	90.5	47	19	AAW77585	Staphylococcus aur
2	19	90.5	81	22	AAU61772	Propionibacterium
3	19	90.5	97	22	AAU41551	Propionibacterium
4	19	90.5	100	22	ABG27830	Novel human diagno
5	19	90.5	113	22	AAU43242	Propionibacterium
6	19	90.5	117	22	AAM84257	Human immune/haema
7	19	90.5	118	20	AAY59973	Human endometrium
8	19	90.5	159	19	AAW31552	Collagen binding p
9	19	90.5	168	22	AAE11855	Staphylococcus aur
10	19	90.5	177	14	AAR39711	A. oryzae WT neutr
11	19	90.5	177	14	AAR39712	A. oryzae C6A neut
12	19	90.5	177	14	AAR39713	A. oryzae C78A neu
13	19	90.5	207	21	AAB24437	Human secreted pro
14	19	90.5	211	19	AAW31553	Collagen binding p
15	19	90.5	225	22	ABG03051	Novel human diagno
16	19	90.5	229	21	AAG05085	Arabidopsis thalia
17	19	90.5	242	22	AAM25472	Human protein sequ
18	19	90.5	243	21	AAG05084	Arabidopsis thalia
19	19	90.5	243	21	AAY95431	Human calcium chan
20	19	90.5	257	21	AAG05083	Arabidopsis thalia
21	19	90.5	264	22	AAU68590	Human novel cytoki
22	19	90.5	270	22	ABB69721	Drosophila melanog
23	19	90.5	270	22	ABG03786	Novel human diagno
24	19	90.5	273	22	AAU31144	Novel human secret
25	19	90.5	290	22	AAU47789	Propionibacterium
26	19	90.5	305	22	ABG02081	Novel human diagno
27	19	90.5	323	22	AAB94964	Human protein sequ
28	19	90.5	330	11	AAR05528	High density lipop
29	19	90.5	332	21	AAY95898	Human myristoylate
30	19	90.5	332	21	AAY95899	Human myristoylate
31	19	90.5	333	21	AAY97409	Zebrafish Hsp-bind
32	19	90.5	341	22	AAG90789	C glutamicum prote
33	19	90.5	341	22	AAB79110	Corynebacterium gl
34	19	90.5	341	22	AAB79144	Corynebacterium gl
35	19	90.5	349	19	AAW44368	Aspergillus nidula
36	19	90.5	352	12	AAR14147	Pre-pro neutral pr
37	19	90.5	363	22	AAU36356	Pseudomonas aerugi
38	19	90.5	366	20	AAY05663	Maize caffeic O-me
39	19	90.5	377	21	AAG20561	Arabidopsis thalia
40	19	90.5	377	21	AAG41800	Arabidopsis thalia
41	19	90.5	386	21	AAG24042	Arabidopsis thalia
42	19	90.5	390	8	AAP70581	Protease biosynthe
43	19	90.5	393	20	AAY35147	Chlamydia pneumoni
44	19	90.5	422	20	AAY28643	Human serine prote
45	19	90.5	422	22	AAB74691	Human protease and
46	19	90.5	427	21	AAG41388	Arabidopsis thalia
47	19	90.5	431	21	AAG41387	Arabidopsis thalia
48	19	90.5	435	21	AAG24041	Arabidopsis thalia
49	19	90.5	440	21	AAG24040	Arabidopsis thalia
50	19	90.5	441	14	AAR31955	Sequence encoded b

ALIGNMENTS

RESULT 1
AAW77585
ID AAW77585 standard; Protein; 47 AA.
XX
AC AAW77585;
XX
DT 30-OCT-1998 (first entry)
XX
DE Staphylococcus aureus protein of unknown function.
XX
KW Staphylococcus aureus protein; immune response induction; eye infection;
KW antibody production; T-cell immune response; gastrointestinal infection;
KW respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW central nervous system; kidney infection; urinary tract infection;
KW antimicrobial compound identification; broad spectrum antibiotic;
KW therapy.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..47
FT /note= "residues designated X are unspecified, and
FT represented as Xaa in the specification"
XX
PN EP841394-A2.
XX
PD 13-MAY-1998.
XX
PF 24-SEP-1997; 97EP-0307485.

XX
PR 24-SEP-1996; 96US-0027032.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Black MT, Burnham MKR, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;
PI Ward JM;
XX
DR WPI; 1998-252940/23.
DR N-PSDB; AAV53383.
XX
PT New nucleic acid sequences from Staphylococcus aureus WCHU29 -
PT useful in vaccines and for treatment of bacterial infections of e.g.
PT respiratory tract and central nervous system
XX
PS Claim 11; Page 267; 390pp; English.
XX
CC This sequence represents a Staphylococcus aureus protein of unknown
CC function, and is encoded by a DNA sequence of the invention.
CC The DNA sequences were isolated from Staphylococcus aureus WCHU29
CC (NCIMB 40771). Host cells containing the DNA sequences are used to
CC produce polypeptides or fragments. The proteins are used in the treatment
CC of disease, for inducing an immune response by administering them, to
CC produce antibody and/or T-cell immune response. Antagonists of the
CC proteins are used for the inhibition of bacterial polypeptides.
CC Conditions which may be treated include bacterial infections, especially
CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
CC urinary tract, skin, bones and joints. The proteins can also be used to
CC identify antimicrobial compounds which are broad spectrum antibiotics,
CC especially useful in the treatment of H. pylori infection.
XX
SQ Sequence 47 AA;

Query Match 90.5%; Score 19; DB 19; Length 47;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 eagxxs 6
          |||  |
Db     31 eagats 36
```

RESULT 2
AAU61772
ID AAU61772 standard; Protein; 81 AA.
XX
AC AAU61772;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #22668.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59620.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 22967; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 81 AA;

Query Match 90.5%; Score 19; DB 22; Length 81;
Best Local Similarity 66.7%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 eagxxs 6
          |||  |
Db     20 eagats 25
```

RESULT 3
AAU41551
ID AAU41551 standard; Protein; 97 AA.
XX
AC AAU41551;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #2447.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX

PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59515.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 2746; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 97 AA;

Query Match 90.5%; Score 19; DB 22; Length 97;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 29 eagsas 34

RESULT 4
ABG27830
ID ABG27830 standard; Protein; 100 AA.
XX
AC ABG27830;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27821.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS92017.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 58189; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 100 AA;

Query Match 90.5%; Score 19; DB 22; Length 100;
Best Local Similarity 66.7%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 95 eagats 100

RESULT 5
AAU43242
ID AAU43242 standard; Protein; 113 AA.
XX
AC AAU43242;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #4138.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX

PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59520.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 4437; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 113 AA;

Query Match 90.5%; Score 19; DB 22; Length 113;
Best Local Similarity 66.7%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 99 eagtas 104

RESULT 6
AAM84257
ID AAM84257 standard; Protein; 117 AA.
XX
AC AAM84257;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:11850.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR N-PSDB; AAK57038.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Claim 11; SEQ ID NO 11850; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 117 AA;

Query Match 90.5%; Score 19; DB 22; Length 117;
Best Local Similarity 66.7%; Pred. No. 7.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 78 eagtas 83

RESULT 7
AAY59973
ID AAY59973 standard; Protein; 118 AA.
XX
AC AAY59973;
XX
DT 31-JAN-2000 (first entry)
XX
DE Human endometrium tumour EST encoded protein 33.
XX
KW Endometrium; human; tumour; cancer; anticancer; cytostatic; EST:
KW treatment; uterine; gene therapy; expressed sequence tag.
XX
OS Homo sapiens.
XX
PN DE19817948-A1.
XX
PD 21-OCT-1999.
XX
PF 17-APR-1998; 98DE-1017948.
XX
PR 17-APR-1998; 98DE-1017948.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-591957/51.
DR N-PSDB; AAZ41991.
XX
PT New nucleic acid sequences expressed in uterine cancer tissues, and
PT derived polypeptides, for treatment of uterine and endometrial cancer
PT and identification of therapeutic agents -
XX
PS Claim 23; Page 288; 444pp; German.
XX
CC This invention describes novel human nucleic acid (cDNA) sequences (A),
CC that are highly expressed in uterine tumour tissue and which have
CC anticancer and cytostatic activity. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of uterine or
CC endometrial cancer; (ii) directly for treating these forms of cancer
CC (including expression from gene therapy vectors) and (iii) for
CC generation of specific antibodies. (A) are identified by assembling ESTs

```
CC   (expressed sequence tags) from a particular tissue type before comparison
CC   of expression patterns. This allows a significantly longer fragment of
CC   the gene to be revealed, so should reduce the number of failures
CC   associated with the fact that ESTs from different libraries may represent
CC   different parts of the same unknown gene, distorting the estimated
CC   frequency of occurrence in a particular tissue. AAY59941-Y60328 represent
CC   protein fragments encoded by the human endometrium tumour cDNA library
CC   derived EST fragments represented in AAZ41981-Z42121.
XX
SQ   Sequence   118 AA;


  Query Match             90.5%;  Score 19;  DB 20;  Length 118;
  Best Local Similarity   66.7%;  Pred. No. 7.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     34 eagsas 39


RESULT   8
AAW31552
ID   AAW31552 standard; Protein; 159 AA.
XX
AC   AAW31552;
XX
DT   21-MAY-1998  (first entry)
XX
DE   Collagen binding protein M17 epitope.
XX
KW   Collagen binding protein; cna gene; sepsis; infection;
KW   microbial surface component regonising adhesive matrix molecule;
KW   MSCRAMM; adhesin; vaccine; immunisation; diagnosis; therapy;
KW   epitope M17.
XX
OS   Staphylcoccus aureus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..12
FT                   /note= "vector pQE30-derived peptide"
FT   Protein         13..159
FT                   /note= "epitope M17"
XX
PN   WO9743314-A2.
XX
PD   20-NOV-1997.
XX
PF   14-MAY-1997;   97WO-US08210.
XX
PR   16-MAY-1996;   96US-0017678.
XX
PA   (UABR-) UAB RES FOUND.
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Hook M,  House-Pompeo K,  Patti JM,  Sthanam N,  Symersky J;
XX
DR   WPI; 1998-008801/01.
DR   N-PSDB; AAT93436.
XX
PT   Antibody that interacts with collagen binding domain of
PT   Staphylococcal cna gene product - useful to prevent bacterial sepsis
PT   in animal infected with Staphylococcus aureus
XX
PS   Claim 31; Page 114; 143pp; English.
XX
CC   This protein comprises Staphylococcus aureus collagen binding
CC   protein (CBP) epitope M17, i.e. amino acids 151-297 of full-length
CC   CBP, plus a vector-derived N-terminal peptide.  Claimed 441, 849
CC   and 1500 bp nucleic acid sequences (see AAT93436-38) respectively
CC   encode CBP epitopes M17, M31 and M55 (see AAW31552-54) that confer
CC   protection against S. aureus infection.  These nucleic acid
CC   sequences can be used in the recombinant production of the CBP
CC   epitopes.  The CBP protein and antigenic epitopes are contemplated
CC   for use in the treatment of pathological infections, especially to
CC   prevent bacterial adhesion to collagen.  The claimed nucleic acids
CC   as well as claimed anti-CBP antibodies will also be of use in
CC   screening, diagnostic and therapeutic applications including active
CC   and passive immunisation and methods for the prevention of
CC   bacterial colonisation in an animal such as a human.  The CBP
CC   epitopes are also contemplated for use in the preparation of
CC   vaccines and as carrier proteins in vaccine formulations, as well
CC   as in the formulation of compositions for the prevention of S.
CC   aureus infection.
XX
SQ   Sequence   159 AA;


  Query Match             90.5%;  Score 19;  DB 19;  Length 159;
  Best Local Similarity   66.7%;  Pred. No. 9.7e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     28 eagtss 33


RESULT   9
AAE11855
ID   AAE11855 standard; Protein; 168 AA.
XX
AC   AAE11855;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Staphylococcus aureus CNA19 protein.
XX
KW   Collagen-binding region; CNA19; Staphylococcus aureus infection;
KW   Staphylococcus epidermidis infection; vaccine; CNA protein;
KW   cross-reactive antibody.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   Region          24..29
FT                   /note= "Beta strand a; this region forms a part of
FT                   the trench in the beta sheet"
FT   Region          37..44
FT                   /note= "Beta strand b; this region forms a part of
FT                   the trench in the beta sheet"
FT   Region          55..61
FT                   /note= "Beta strand c"
FT   Region          65..78
FT                   /note= "Beta strand d; a portion of this region forms
FT                   a part of the trench in the beta sheet"
FT   Region          82..84
FT                   /note= "Beta strand e; a portion of this region forms
FT                   a part of the trench in the beta sheet"
FT   Region          89..92
FT                   /note= "Alpha helix 1"
FT   Region          93..96
FT                   /note= "Alpha helix 2"
FT   Region          101..105
FT                   /note= "Beta strand f"
FT   Region          110..115
FT                   /note= "Beta strand g"
FT   Region          123..133
FT                   /note= "Beta strand h; a portion of this region forms
FT                   a part of the trench in the beta sheet"
FT   Region          140..149
FT                   /note= "Beta strand i"
FT   Region          157..167
FT                   /note= "Beta strand j"
XX
PN   WO200170267-A1.
XX
```

```
PD   27-SEP-2001.
XX
PF   19-MAR-2001; 2001WO-US08554.
XX
PR   17-MAR-2000; 2000US-189968P.
PR   25-APR-2000; 2000US-199370P.
PR   15-AUG-2000; 2000US-225402P.
XX
PA   (INHI-) INHIBITEX INC.
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
PA   (UYPA-) UNIV PAVIA.
XX
PI   Hook M,  Xu Y,  Speziale P,  Visai L,  Casolini F,  Patti J,  Patel P;
PI   Domanski P;
XX
DR   WPI; 2001-607512/69.
XX
PT   Novel isolated antibody which recognizes collagen-binding peptide such
PT   as CNA19 peptide from Staphylococcus aureus, useful for preventing or
PT   treating Staphylococcus aureus or epidermidis infection -
XX
PS   Example 2; Fig 2A; 107pp; English.
XX
CC   The invention relates to an antibody which recognises a collagen-binding
CC   region including CNA19 of CNA protein from Staphylococcus aureus. This
CC   antibody is cross-reactive to collagen binding region of both S. aureus
CC   and S. epidermidis. It is useful for preventing or treating S. aureus or
CC   S. epidermidis infection in human or animal, and for displacing S. aureus
CC   or S. epidermidis bound to collagen. Antibody of the invention is useful
CC   for interfering with, modulating, and inhibiting the binding interactions
CC   between Staphylococcal bacteria and collagen, for detecting the presence
CC   of Staphylococcal bacteria or Staphylococcal collagen or binding
CC   proteins, to diagnose Staphylococcal infection, as research tools, for
CC   development of vaccine for passive immunisation against Staphylococcal
CC   infections, and in production facilities or laboratories to isolate
CC   additional quantities of collagen-binding proteins. It is also useful
CC   for treating medical instruments in order to reduce or eliminate the
CC   possibility of them becoming infected or further spreading the
CC   infection. The present sequence is S. aureus CNA19 protein.
XX
SQ   Sequence   168 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 168;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     16 eagtss 21


RESULT  10
AAR39711
ID   AAR39711 standard; Protein; 177 AA.
XX
AC   AAR39711;
XX
DT   24-JAN-1994  (first entry)
XX
DE   A. oryzae WT neutral protease.
XX
KW   Wildtype; protease; varient; neutral; mercapto group; food products;
KW   thermal stability; soy fermantation.
XX
OS   Aspergillus oryzae.
XX
PN   JP05168479-A.
XX
PD   02-JUL-1993.
XX
PF   26-DEC-1991;   91JP-0344443.
XX
PR   26-DEC-1991;   91JP-0344443.
XX
PA   (SHSA ) SHOKUHIN SANGYO KOSOKINO HENKA.
XX
DR   WPI; 1993-247571/31.
DR   N-PSDB; AAQ46955.
XX
PT   New variant neutral protease II - includes cysteine substd. with
PT   aminoacid having no mercapto gp. in aminoacid sequence of yellow
PT   green koji mould neutral protease II
XX
PS   Disclosure; Page 7-8; 9pp; Japanese.
XX
CC   This sequence represents a wildtype protease which may be used as the
CC   basis for the production of a varient neutral protease.  The varient
CC   protease has either Cys6 or Cys78 substituted with an amino acid
CC   which has no -SH group.  The varient proteases (see also AAR39713-14)
CC   have lower thermal stability than the WT and may be used in soy
CC   fermantation microorganisms.  Soy produced by these microorganisms
CC   may be made into food products which will not be degraded by the
CC   presence of protease.
XX
SQ   Sequence   177 AA;


  Query Match             90.5%;  Score 19;  DB 14;  Length 177;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     65 eagsts 70


RESULT  11
AAR39712
ID   AAR39712 standard; Protein; 177 AA.
XX
AC   AAR39712;
XX
DT   24-JAN-1994  (first entry)
XX
DE   A. oryzae C6A neutral protease.
XX
KW   Wildtype; protease; varient; neutral; mercapto group; food products;
KW   thermal stability; soy fermantation.
XX
OS   Aspergillus oryzae.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /label= C6A
XX
PN   JP05168479-A.
XX
PD   02-JUL-1993.
XX
PF   26-DEC-1991;   91JP-0344443.
XX
PR   26-DEC-1991;   91JP-0344443.
XX
PA   (SHSA ) SHOKUHIN SANGYO KOSOKINO HENKA.
XX
DR   WPI; 1993-247571/31.
XX
PT   New variant neutral protease II - includes cysteine substd. with
PT   aminoacid having no mercapto gp. in aminoacid sequence of yellow
PT   green koji mould neutral protease II
XX
PS   Disclosure; Page 8; 9pp; Japanese.
XX
CC   The sequences given in AAR39712-13 represent varient neutral proteases
CC   based on the Aspergillus oryzae protease sequence (see also AAR39711).
```

```
CC   These varient proteases have either Cys6 or Cys78 substituted with an
CC   amino acid which has no -SH group.  These varient proteases have lower
CC   thermal stability than the WT and may be used in soy fermantation
CC   microorganisms.  Soy produced by these microorganisms may be made into
CC   food products which will not be degraded by the presence of protease.
XX
SQ   Sequence   177 AA;


  Query Match            90.5%;  Score 19;  DB 14;  Length 177;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     65 eagsts 70


RESULT  12
AAR39713
ID   AAR39713 standard; Protein; 177 AA.
XX
AC   AAR39713;
XX
DT   24-JAN-1994  (first entry)
XX
DE   A. oryzae C78A neutral protease.
XX
KW   Wildtype; protease; varient; neutral; mercapto group; food products;
KW   thermal stability; soy fermantation.
XX
OS   Aspergillus oryzae.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 78
FT                   /label= C78A
XX
PN   JP05168479-A.
XX
PD   02-JUL-1993.
XX
PF   26-DEC-1991;   91JP-0344443.
XX
PR   26-DEC-1991;   91JP-0344443.
XX
PA   (SHSA ) SHOKUHIN SANGYO KOSOKINO HENKA.
XX
DR   WPI; 1993-247571/31.
XX
PT   New variant neutral protease II - includes cysteine substd. with
PT   aminoacid having no mercapto gp. in aminoacid sequence of yellow
PT   green koji mould neutral protease II
XX
PS   Disclosure; Page 8; 9pp; Japanese.
XX
CC   The sequences given in AAR39712-13 represent varient neutral proteases
CC   based on the Aspergillus oryzae protease sequence (see also AAR39711).
CC   These varient proteases have either Cys6 or Cys78 substituted with an
CC   amino acid which has no -SH group.  These varient proteases have lower
CC   thermal stability than the WT and may be used in soy fermantation
CC   microorganisms.  Soy produced by these microorganisms may be made into
CC   food products which will not be degraded by the presence of protease.
XX
SQ   Sequence   177 AA;


  Query Match            90.5%;  Score 19;  DB 14;  Length 177;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     65 eagsts 70


RESULT  13
AAB24437
ID   AAB24437 standard; Protein; 207 AA.
XX
AC   AAB24437;
XX
DT   20-NOV-2000  (first entry)
XX
DE   Human secreted protein sequence encoded by gene 1 SEQ ID NO:62.
XX
KW   Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW   antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW   antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
KW   neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW   immune system disorder; angiogenesis; hyperproliferative disorder;
KW   cardiovascular disorder; apoptosis; neurological disease;
KW   infectious disease; wound healing.
XX
OS   Homo sapiens.
XX
PN   WO200035937-A1.
XX
PD   22-JUN-2000.
XX
PF   16-DEC-1999;   99WO-US29950.
XX
PR   17-DEC-1998;   98US-0112809.
PR   18-DEC-1998;   98US-0113006.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R,  Rosen CA,  Endress GA,  Soppet DR,  Ni J;
PI   Duan DR,  Moore PA,  Shi Y,  Lafleur DW,  Olsen HS,  Florence K;
XX
DR   WPI; 2000-431566/37.
DR   N-PSDB; AAA78381.
XX
PT   Forty seven human nucleic acids encoding secreted proteins, useful in
PT   the treatment, prevention and diagnosis of cancers, disorders of the
PT   immune system, angiogenesis disorders, neurological diseases and
PT   hyperproliferative disorders -
XX
PS   Claim 11; Page 478-479; 562pp; English.
XX
CC   The polynucleotide sequence given in AAA78381 to AAA78432 encode the
CC   human secreted proteins given in AAB24437 to AAB24604. Human secreted
CC   proteins have activities based on the tissues and cells the genes are
CC   expressed in. Examples of activities include: cytostatic; antianaemic;
CC   antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
CC   antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
CC   nootropic; neuroprotective; antimicrobial and antiparkinsonian.
CC   Human secreted protein polynucleotides, polypeptides, antagonists and/or
CC   agonists may be useful in treating, preventing, and/or diagnosing other
CC   diseases, disorders, and/or conditions such as: (a) cancers; (b)
CC   disorders of the immune system; (c) angiogenesis disorders; (d)
CC   hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
CC   associated with increase apoptosis; (g) neurological diseases; and
CC   (h) infectious diseases. They are also used to promote wound healing.
CC   AAA78372 to AAA78380 and AAB24436 represent sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   207 AA;


  Query Match            90.5%;  Score 19;  DB 21;  Length 207;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     77 eagats 82
```

```
RESULT  14
AAW31553
ID   AAW31553 standard; Protein; 211 AA.
XX
AC   AAW31553;
XX
DT   21-MAY-1998  (first entry)
XX
DE   Collagen binding protein M31 epitope.
XX
KW   Collagen binding protein; cna gene; sepsis; infection;
KW   microbial surface component regonising adhesive matrix molecule;
KW   MSCRAMM; adhesin; vaccine; immunisation; diagnosis; therapy;
KW   epitope M31.
XX
OS   Staphylcoccus aureus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..12
FT                   /note= "vector pQE30-derived peptide"
FT   Protein         13..211
FT                   /note= "epitope M31"
XX
PN   WO9743314-A2.
XX
PD   20-NOV-1997.
XX
PF   14-MAY-1997;   97WO-US08210.
XX
PR   16-MAY-1996;   96US-0017678.
XX
PA   (UABR-) UAB RES FOUND.
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Hook M,  House-Pompeo K,  Patti JM,  Sthanam N,  Symersky J;
XX
DR   WPI; 1998-008801/01.
DR   N-PSDB; AAT93437.
XX
PT   Antibody that interacts with collagen binding domain of
PT   Staphylococcal cna gene product - useful to prevent bacterial sepsis
PT   in animal infected with Staphylococcus aureus
XX
PS   Claim 31; Page 115-116; 143pp; English.
XX
CC   This protein comprises Staphylococcus aureus collagen binding
CC   protein (CBP) epitope M31, i.e. amino acids 61-343 of full-length
CC   CBP, plus a vector-derived N-terminal peptide.  Claimed 441, 849
CC   and 1500 bp nucleic acid sequences (see AAT93436-38) respectively
CC   encode CBP epitopes M17, M31 and M55 (see AAW31552-54) that confer
CC   protection against S. aureus infection.  These nucleic acid
CC   sequences can be used in the recombinant production of the CBP
CC   epitopes.  The CBP protein and antigenic epitopes are contemplated
CC   for use in the treatment of pathological infections, especially to
CC   prevent bacterial adhesion to collagen.  The claimed nucleic acids
CC   as well as claimed anti-CBP antibodies will also be of use in
CC   screening, diagnostic and therapeutic applications including active
CC   and passive immunisation and methods for the prevention of
CC   bacterial colonisation in an animal such as a human.  The CBP
CC   epitopes are also contemplated for use in the preparation of
CC   vaccines and as carrier proteins in vaccine formulations, as well
CC   as in the formulation of compositions for the prevention of S.
CC   aureus infection.
XX
SQ   Sequence   211 AA;

  Query Match           90.5%;  Score 19;  DB 19;  Length 211;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      34 eagtss 39


RESULT  15
ABG03051
ID   ABG03051 standard; Protein; 225 AA.
XX
AC   ABG03051;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #3042.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS67238.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 33410; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   225 AA;

  Query Match           90.5%;  Score 19;  DB 22;  Length 225;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
```

```
          ||| '|
Db     16 eagtss 21


RESULT  16
AAG05085
ID   AAG05085 standard; Protein; 229 AA.
XX
AC   AAG05085;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1356.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
```

```
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             90.5%;  Score 19;  DB 21;  Length 229;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     156 eagsss 161


RESULT  17
AAM25472
ID   AAM25472 standard; Protein; 242 AA.
XX
AC   AAM25472;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:987.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35017.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   N-PSDB; AAH99413.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS   Claim 20; Page 207; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and
CC   cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production, The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders.
XX
SQ   Sequence   242 AA;

  Query Match             90.5%;  Score 19;  DB 22;  Length 242;
```

```
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     203 eagsss 208


RESULT  18
AAG05084
ID   AAG05084 standard; Protein; 243 AA.
XX
AC   AAG05084;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1355.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
```

```
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             90.5%;  Score 19;  DB 21;  Length 243;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     170 eagsss 175
```

```
RESULT  19
AAY95431
ID   AAY95431 standard; Protein; 243 AA.
XX
AC   AAY95431;
XX
DT   10-OCT-2000  (first entry)
XX
DE   Human calcium channel SOC-2/CRAC-1.
XX
KW   SOC-2/CRAC-1; calcium channel; human; store operated channel;
KW   calcium release activated channel; therapy; diagnosis;
KW   lymphocyte proliferative disorder.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /note= "encoded by SCA"
FT   Misc-difference 104
FT                   /note= "encoded by CSA"
FT   Misc-difference 105
FT                   /note= "encoded by RSC"
FT   Misc-difference 109
FT                   /note= "encoded by GNT"
FT   Misc-difference 141
FT                   /note= "encoded by NCA"
XX
PN   WO200040614-A2.
XX
PD   13-JUL-2000.
XX
PF   20-DEC-1999;   99WO-US29996.
XX
PR   30-DEC-1998;   98US-0114220.
PR   29-JAN-1999;   99US-0120018.
PR   22-JUN-1999;   99US-0140415.
XX
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI   Scharenberg AM;
XX
DR   WPI; 2000-465957/40.
DR   N-PSDB; AAA49918.
XX
PT   New SOC/CRAC calcium channel polynucleotides and polypeptides used to
PT   diagnose and treat proliferative disorders associated with the channel,
PT   and to screen for novel modulators of the channel -
XX
PS   Claim 14; Page 58-59; 108pp; English.
XX
CC   The present sequence is that of a partial sequence of human
CC   SOC-2/CRAC-1 (full-length sequence given in AAY95435), as deduced
CC   from a partial cDNA clone (see AAA49918).  SOC-2/CRAC-1 is a member
CC   of a novel family of store operated channel (SOC) or calcium release
CC   activated channel (CRAC) polypeptides that modulate Ca2+ flux into
CC   and out of a cell, and which may be activated upon depletion of
CC   Ca2+ from intracellular calcium stores, allowing Ca2+ influx into
CC   a cell.  SOC-2/CRAC-1 is expressed predominantly in human
CC   haematopoietic cells, liver, spleen, heart and kidney.
CC   Compositions for expressing functional SOC/CRAC calcium channel
CC   polypeptides in cells are expected to be useful for treating
CC   patients that have reduced extracellular calcium influx into their
CC   SOC/CRAC-expressing cells.  They will also be useful for delivering
CC   therapeutic and/or imaging agents to such cells to modulate
CC   proliferation and growth.  SOC/CRAC polypeptides also represent
CC   targets for designing and/or identifying inhibitors that block
CC   lymphocyte proliferation and binding agents that selectively bind
CC   to SOC/CRAC polypeptides to which drugs or toxins can be conjugated
CC   for delivery to SOC/CRAC expressing cells.  Methods for determining
CC   the level of SOC/CRAC expression in a subject can be used to assess
CC   the presence, or absence, or stage of a proliferative disorder,
```

```
CC   e.g. a lymphocyte proliferative disorder.
XX
SQ   Sequence   243 AA;


  Query Match             90.5%;  Score 19;  DB 21;  Length 243;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
             |||  |
Db      53 eagsss 58


RESULT  20
AAG05083
ID   AAG05083 standard; Protein; 257 AA.
XX
AC   AAG05083;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1354.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
```

```
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           90.5%;  Score 19;  DB 21;  Length 257;
  Best Local Similarity 66.7%;  Pred. No. 1.6e+03;

    Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     184 eagsss 189


RESULT  21
AAU68590
ID   AAU68590 standard; Protein; 264 AA.
XX
AC   AAU68590;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human novel cytokine encoded by cDNA 790CIP2D_8 #2.
XX
KW   Human; cytokine; cell proliferation; cell differentiation;
KW   antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
KW   nervous system disease; neuropathy; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; spinal cord disorder;
KW   head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
KW   platelet disorder; thrombocytopaenia; stem cell disorder;
KW   aplastic anaemia; tissue regeneration; wound healing; ulcer;
KW   osteoporosis; osteoarthritis; bone degenerative disorder;
KW   periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
KW   severe combined immunodeficiency; infection; autoimmune disorder;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
KW   asthma; coagulation disorder; haemophilia; sepsis; nephritis;
KW   inflammatory bowel disease; food supplement; immunogen.
XX
OS   Homo sapiens.
XX
PN   WO200175093-A1.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US10484.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
PR   22-SEP-2000; 2000US-0668680.
PR   23-OCT-2000; 2000US-0695618.
PR   30-NOV-2000; 2000US-0728711.
PR   14-MAR-2001; 2000US-0728711.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Asundi V,  Zhou P,  Xue AJ,  Ren F,  Zhang J,  Wang J,  Xu C;
PI   Yang Y,  Zaho QA,  Chen R,  Wang D,  Goodrich RW,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-626432/72.
DR   N-PSDB; AAS59882.
XX
PT   New polypeptides and nucleic acids, useful for diagnosis, treatment of
PT   inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
PT   degenerative disorders, cancer and promoting wound healing  -
XX
PS   Claim 20; Page 328; 336pp; English.
XX
CC   The invention relates to isolated human polypeptides (which may be
CC   cytokines) and the polynucleotides encoding them. The protein is useful
CC   for identifying a compound which binds to it (e.g. modulators, agonists
CC   and antagonists). The polynucleotides are useful as an array for mismatch
CC   detection. The proteins and nucleic acids are useful as nutritional
CC   sources or supplements. The protein exhibits exhibits activity relating
CC   to cytokine, cell proliferation, cell differentiation, antiinflammatory,
CC   stem cell growth factor activity, immune stimulating or immune
CC   suppressing and activin or inhibin related activities. The proteins (and
CC   antibodies raised against them) and nucleic acids are therefore useful in
CC   the diagnosis and treatment of diseases and disorders such as cancer,
CC   central and peripheral nervous system diseases and neuropathies,
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
```

CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, and in tissue repair, healing of burns, incisions, ulcers, for
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
CC periodontal disease, lung or liver fibrosis, reperfusion injury in
CC various tissues, various immune deficiencies and disorders including
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
CC such as asthma or other respiratory problems, coagulation disorders,
CC haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory
CC bowel disease, viral infection and are useful in altering bodily
CC characteristics. The present sequence represents a novel protein of the
CC invention.
XX
SQ Sequence 264 AA;

Query Match 90.5%; Score 19; DB 22; Length 264;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 207 eagaas 212

RESULT 22
ABB69721
ID ABB69721 standard; Protein; 270 AA.
XX
AC ABB69721;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 35955.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL13824.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 35955; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 270 AA;

Query Match 90.5%; Score 19; DB 22; Length 270;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 169 eagsts 174

RESULT 23
ABG03786
ID ABG03786 standard; Protein; 270 AA.
XX
AC ABG03786;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3777.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS67973.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 34145; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

```
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   270 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 270;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     16 eagtss 21


RESULT  24
AAU31144
ID   AAU31144 standard; Protein; 273 AA.
XX
AC   AAU31144;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #1635.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 409; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   273 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 273;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     35 eagass 40


RESULT  25
AAU47789
ID   AAU47789 standard; Protein; 290 AA.
XX
AC   AAU47789;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #8685.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
DR   N-PSDB; AAS59539.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Example 1; SEQ ID No 8984; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   290 AA;
```

```
  Query Match             90.5%;  Score 19;  DB 22;  Length 290;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     79 eagass 84


RESULT  26
ABG02081
ID   ABG02081 standard; Protein; 305 AA.
XX
AC   ABG02081;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #2072.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS66268.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 32440; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   305 AA;
```

```
  Query Match             90.5%;  Score 19;  DB 22;  Length 305;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    178 eagaas 183


RESULT  27
AAB94964
ID   AAB94964 standard; Protein; 323 AA.
XX
AC   AAB94964;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:16523.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 16523; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   323 AA;
```

```
  Query Match          90.5%;  Score 19;  DB 22;  Length 323;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    308 eagsts 313


RESULT  28
AAR05528
ID   AAR05528 standard; protein; 330 AA.
XX
AC   AAR05528;
XX
DT   23-OCT-1990  (first entry)
XX
DE   High density lipoprotein (HDL) binding protein.
XX
KW   High density lipoprotein; HDL-binding protein; atherosclerosis;
KW   hypercholesterolaemia; ds.
XX
OS   Homo sapiens.
XX
PN   WO9005744-A.
XX
PD   31-MAY-1990.
XX
PF   17-NOV-1989;   89WO-0005169.
XX
PR   18-NOV-1988;   88US-0273388.
XX
PA   (UNIW ) UNIV OF WASHINGTON.
PA   (ZYMO-) ZYMOGENETICS INC.
XX
PI   Oram JF,  McKnight GL,  Hart CE,  Curtis DA;
XX
DR   WPI; 1990-193405/25.
DR   N-PSDB; AAQ04784.
XX
PT   New mammalian proteins binding high density lipoprotein sub-class 3 -
PT   DNA encoding them and derived antibodies, for screening
PT   potentially therapuetic HDL analogues and for diagnosing risk of
PT   atherosclerosis.
XX
PS   Claim 4; Fig 1A-D; 79pp; English.
XX
CC   The protein product may be used to raise Abs, and the cDNA to
CC   create probes, both useful in screening for HDL analogues,
CC   agonists and antagonists, and in identifying abnormalities in the
CC   HDL binding/receptor pathway. HDL analogues can be used in treating
CC   hypercholestrolaemia and atherosclerosis
XX
SQ   Sequence   330 AA;


  Query Match          90.5%;  Score 19;  DB 11;  Length 330;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    205 eagaas 210


RESULT  29
AAY95898
ID   AAY95898 standard; Protein; 332 AA.
XX
AC   AAY95898;
XX
DT   20-NOV-2000  (first entry)
XX
DE   Human myristoylated alanine-rich C kinase substrate MARCKS.
XX
KW   MARCKS; myristoylated alanine-rich C kinase substrate; human;
KW   mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
KW   chronic obstructive pulmonary disease; pneumonia; empysema;
KW   influenza; rhinitis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 84
FT                   /note= "Ser in sequence of AAY95899"
FT   Misc-difference 119
FT                   /note= "Ala in sequence of AAY95899"
FT   Peptide         2..25
FT                   /note= "MANS peptide of AAY95896"
FT   Peptide         152..176
FT                   /note= "MA-PSD peptide of AAY95897"
XX
PN   WO200050062-A2.
XX
PD   31-AUG-2000.
XX
PF   24-FEB-2000; 2000WO-US05050.
XX
PR   24-FEB-1999;   99US-0256154.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Li Y,  Martin LD,  Adler KB;
XX
DR   WPI; 2000-572036/53.
DR   N-PSDB; AAA50339.
XX
PT   Regulating mucus secretion by a mucus-secreting cell, useful for
PT   treating e.g. bronchitis, asthma or pneumonia, by administering a
PT   compound that inhibits or enhances myristolated alanine-rich C-kinase
PT   substrate protein -
XX
PS   Claim 3; Page 42-43; 66pp; English.
XX
CC   The present sequence is that of human myristoylated alanine-rich C
CC   kinase substrate MARCKS protein, a major cellular substrate.  The
CC   invention relates to methods of inhibiting mucus secretion by a
CC   mucus-secreting cell by administering a compound that inhibits
CC   MARCKS protein-related mucus secretion.  Such compounds include
CC   active fragments of MARCKS protein such as MANS peptide (see
CC   AAY95897) and MA-PDS peptide (see AAY95897), which corresponds to a
CC   phosphorylation site of MARCKS.  The inhibitor compounds can be
CC   used to treat conditions such as bronchitis, cystic fibrosis,
CC   chronic obstructive pulmonary disease, asthma, emphysema,
CC   pneumonia, influenza, rhinitis and the common cold.  An alternative
CC   sequence for MARCKS is provided in AAY95899, which differs from the
CC   present sequence at 2 amino acid residues, Ala-84 (Ser) and
CC   Pro-119 (Ala).
XX
SQ   Sequence   332 AA;


  Query Match          90.5%;  Score 19;  DB 21;  Length 332;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    206 eagaas 211


RESULT  30
AAY95899
ID   AAY95899 standard; Protein; 332 AA.
```

```
XX
AC   AAY95899;
XX
DT   20-NOV-2000  (first entry)
XX
DE   Human myristoylated alanine-rich C kinase substrate MARCKS.
XX
KW   MARCKS; myristoylated alanine-rich C kinase substrate; human;
KW   mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
KW   chronic obstructive pulmonary disease; pneumonia; empysema;
KW   influenza; rhinitis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 84
FT                   /note= "Ala in sequence of AAY95898"
FT   Misc-difference 119
FT                   /note= "Pro in sequence of AAY95898"
FT   Peptide         2..25
FT                   /note= "MANS peptide of AAY95896"
FT   Peptide         152..176
FT                   /note= "MA-PSD peptide of AAY95897"
XX
PN   WO200050062-A2.
XX
PD   31-AUG-2000.
XX
PF   24-FEB-2000; 2000WO-US05050.
XX
PR   24-FEB-1999;   99US-0256154.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Li Y,  Martin LD,  Adler KB;
XX
DR   WPI; 2000-572036/53.
DR   N-PSDB; AAA50340.
XX
PT   Regulating mucus secretion by a mucus-secreting cell, useful for
PT   treating e.g. bronchitis, asthma or pneumonia, by administering a
PT   compound that inhibits or enhances myristolated alanine-rich C-kinase
PT   substrate protein -
XX
PS   Disclosure; Page 46-47; 66pp; English.
XX
CC   The present sequence is that of human myristoylated alanine-rich C
CC   kinase substrate MARCKS protein, a major cellular substrate.  The
CC   invention relates to methods of inhibiting mucus secretion by a
CC   mucus-secreting cell by administering a compound that inhibits
CC   MARCKS protein-related mucus secretion.  Such compounds include
CC   active fragments of MARCKS protein such as MANS peptide (see
CC   AAY95897) and MA-PDS peptide (see AAY95897), which corresponds to a
CC   phosphorylation site of MARCKS.  The inhibitor compounds can be
CC   used to treat conditions such as bronchitis, cystic fibrosis,
CC   chronic obstructive pulmonary disease, asthma, emphysema,
CC   pneumonia, influenza, rhinitis and the common cold.  An alternative
CC   sequence for MARCKS is provided in AAY95898, which differs from the
CC   present sequence at 2 amino acid residues, Ser-84 (Ala) and
CC   Ala-119 (Pro).
XX
SQ   Sequence   332 AA;


  Query Match             90.5%;  Score 19;  DB 21;  Length 332;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    206 eagaas 211
```

```
RESULT  31
AAY97409
ID   AAY97409 standard; Protein; 333 AA.
XX
AC   AAY97409;
XX
DT   14-SEP-2000  (first entry)
XX
DE   Zebrafish Hsp-binding protein HspBPF.
XX
KW   Zebrafish; heat shock protein binding protein; HspBPF;
KW   development; apoptosis; cellular stress; heart disease; cancer;
KW   hypoxia.
XX
OS   Brachydanio rerio.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 81
FT                   /label= unknown
FT                   /note= "encoded by GAA"
FT   Misc-difference 90
FT                   /label= unknown
FT                   /note= "encoded by GAC"
FT   Misc-difference 91
FT                   /label= unknown
FT                   /note= "encoded by GAG"
XX
PN   WO200031114-A1.
XX
PD   02-JUN-2000.
XX
PF   19-NOV-1999;   99WO-US27651.
XX
PR   20-NOV-1998;   98US-0109351.
XX
PA   (GUER/) GUERRIERO V.
PA   (RAYN/) RAYNES D A.
XX
PI   Guerriero V,  Raynes DA;
XX
DR   WPI; 2000-400030/34.
DR   N-PSDB; AAA38747.
XX
PT   New polynucleotides encoding human heat-shock protein-binding protein,
PT   HspBP-1 and HspBP-2, useful for investigating the effects of heat
PT   shock-protein regulation -
XX
PS   Disclosure; Page 35-36; 38pp; English.
XX
CC   The present sequence is the sequence for the zebrafish heat shock
CC   protein-binding protein (HspBPF), which is involved in the
CC   regulation of the heat shock proteins, which are components of the
CC   development, apoptosis and cellular stress pathways. The human
CC   homologue is useful not only for research into these areas, but also
CC   for treating disorders such as heart disease, hypoxia and cancer.
XX
SQ   Sequence   333 AA;


  Query Match             90.5%;  Score 19;  DB 21;  Length 333;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     23 eagsas 28


RESULT  32
AAG90789
ID   AAG90789 standard; Protein; 341 AA.
XX
AC   AAG90789;
```

XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 4543.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH66008.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT -
XX
PS Claim 17; SEQ ID NO: 4543; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 341 AA;

Query Match 90.5%; Score 19; DB 22; Length 341;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 eagxxs 6
          |||  |
Db    239 eagtss 244
```

RESULT 33
AAB79110
ID AAB79110 standard; Protein; 341 AA.
XX
AC AAB79110;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum HA protein sequence SEQ ID NO:176.
XX
KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; genetic engineering;
KW Brevibacterium; environmental condition.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100842-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00911.
XX
PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032127.
PR 09-JUL-1999; 99DE-1032128.
PR 09-JUL-1999; 99DE-1032129.
PR 09-JUL-1999; 99DE-1032226.
PR 14-JUL-1999; 99DE-1032920.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1032930.
PR 14-JUL-1999; 99DE-1032933.
PR 14-JUL-1999; 99DE-1032935.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033002.
PR 14-JUL-1999; 99DE-1033003.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041390.
PR 31-AUG-1999; 99DE-1041391.
PR 03-SEP-1999; 99DE-1042088.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-061974/07.
DR N-PSDB; AAF71225.
XX
PT New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
PT or enzymes -
XX
PS Claim 20; Page 383-384; 712pp; English.
XX
CC AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
CC C. glutamicum HA genes (I) can be used in vectors for expression in host
CC cells and production of fine chemicals, such as, an organic acid,
CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (I) or HA proteins encoded by then are
CC used for diagnosing the presence or activity of Corynebacterium
CC diphtheriae. (I) can be used to map the C. glutamicum genome or can be
CC used as markers for genetically engineered Corynebacterium or
CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain
CC homeostasis in C. glutamicum or help the microorganism to adapt to
CC different environmental conditions.
XX

```
SQ   Sequence   341 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 341;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    239 eagtss 244


RESULT  34
AAB79144
ID   AAB79144 standard; Protein; 341 AA.
XX
AC   AAB79144;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum HA protein sequence SEQ ID NO:244.
XX
KW   Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; genetic engineering;
KW   Brevibacterium; environmental condition.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100842-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00911.
XX
PR   25-JUN-1999;   99US-0141031.
PR   08-JUL-1999;   99DE-1031636.
PR   09-JUL-1999;   99DE-1032125.
PR   09-JUL-1999;   99DE-1032126.
PR   09-JUL-1999;   99DE-1032127.
PR   09-JUL-1999;   99DE-1032128.
PR   09-JUL-1999;   99DE-1032129.
PR   09-JUL-1999;   99DE-1032226.
PR   14-JUL-1999;   99DE-1032920.
PR   14-JUL-1999;   99DE-1032922.
PR   14-JUL-1999;   99DE-1032924.
PR   14-JUL-1999;   99DE-1032928.
PR   14-JUL-1999;   99DE-1032930.
PR   14-JUL-1999;   99DE-1032933.
PR   14-JUL-1999;   99DE-1032935.
PR   14-JUL-1999;   99DE-1032973.
PR   14-JUL-1999;   99DE-1033002.
PR   14-JUL-1999;   99DE-1033003.
PR   14-JUL-1999;   99DE-1033005.
PR   14-JUL-1999;   99DE-1033006.
PR   31-AUG-1999;   99DE-1041378.
PR   31-AUG-1999;   99DE-1041379.
PR   31-AUG-1999;   99DE-1041390.
PR   31-AUG-1999;   99DE-1041391.
PR   03-SEP-1999;   99DE-1042088.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M,  Kroeger B,  Schroeder H,  Zelder O,  Haberhauer G;
XX
DR   WPI; 2001-061974/07.
DR   N-PSDB; AAF71259.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid for production or
PT   modulation of production of fine chemicals such as amino acids,
PT   nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
PT   or enzymes -
XX
PS   Claim 20; Page 466-467; 712pp; English.
XX
CC   AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
CC   and adaptation (HA) proteins given in AAB79023 to AAB79242. The
CC   C. glutamicum HA genes (I) can be used in vectors for expression in host
CC   cells and production of fine chemicals, such as, an organic acid,
CC   proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC   pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC   fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC   polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC   glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC   cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC   tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC   be modulated. The presence of (I) or HA proteins encoded by then are
CC   used for diagnosing the presence or activity of Corynebacterium
CC   diphtheriae. (I) can be used to map the C. glutamicum genome or can be
CC   used as markers for genetically engineered Corynebacterium or
CC   Brevibacterium. The HA proteins encoded by the (I) are used to maintain
CC   homeostasis in C. glutamicum or help the microorganism to adapt to
CC   different environmental conditions.
XX
SQ   Sequence   341 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 341;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    239 eagtss 244


RESULT  35
AAW44368
ID   AAW44368 standard; Protein; 349 AA.
XX
AC   AAW44368;
XX
DT   29-MAY-1998  (first entry)
XX
DE   Aspergillus nidulans metallo-protease pepH.
XX
KW   Aspergillus nidulans; metallo-protease; pepH; protein degradation;
KW   fungus; food processing.
XX
OS   Aspergillus nidulans.
XX
PN   WO9746689-A1.
XX
PD   11-DEC-1997.
XX
PF   05-JUN-1997;   97WO-EP02982.
XX
PR   05-JUN-1996;   96EP-0201579.
XX
PA   (KONN ) GIST-BROCADES BV.
XX
PI   Van Den Hombergh JPTW,  Visser J;
XX
DR   WPI; 1998-042197/04.
DR   N-PSDB; AAV15305, AAV15306.
XX
PT   Metallo-protease deficient fungus with site selected DNA disruption
PT   - and Aspergillus metallo-protease genes, useful in protein
PT   production to reduce protease activity hence protein degradation
XX
PS   Example 4; Page 26-27; 53pp; English.
XX
CC   The present sequence represents a metallo-protease, pepH, from
```

```
CC   Aspergillus nidulans from the present invention. The present invention
CC   describes a new protease deficient filamentous fungus, optionally with
CC   reduced extracellular acid protease activity, containing a site selected
CC   disruption of DNA resulting in reduced metallo-protease activity. The
CC   fungi are useful for the production of (heterologous and homologous)
CC   proteins e.g. for food processing, since reduced protease activity
CC   minimises the chance that, and rate at which, the proteins are degraded
CC   during production. DNA sequences encoding metallo-proteases can be
CC   used to produce metallo-protease deficient fungi, by transforming a
CC   filamentous fungus mutant with the constructs and selecting a
CC   transformed fungus with reduced metallo-protease activity. They are also
CC   useful for producing filamentous fungal metallo-protease, by culturing
CC   filamentous fungi transformed with the constructs under suitable
CC   conditions for sequence expression and recovering the metallo-protease.
CC   Such metallo-proteases are useful to assess in vitro whether proteins
CC   which it is proposed to produce from a fungal host are susceptible to
CC   the protease, so determining which metallo-protease genes need to be
CC   inactivated in the host. They are also useful in industrial processes.
XX
SQ   Sequence   349 AA;


  Query Match             90.5%;  Score 19;  DB 19;  Length 349;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    237 eagsts 242


RESULT  36
AAR14147
ID   AAR14147 standard; Protein; 352 AA.
XX
AC   AAR14147;
XX
DT   12-DEC-1991  (first entry)
XX
DE   Pre-pro neutral protease II.
XX
KW   recombinant enzyme.
XX
OS   Aspergillus oryzae.
XX
FH   Key             Location/Qualifiers
FT   Peptide         176..352
FT                   /note= "mature neutral protease II"
XX
PN   JP03198779-A.
XX
PD   29-AUG-1991.
XX
PF   27-DEC-1989;   89JP-0336737.
XX
PR   27-DEC-1989;   89JP-0336737.
XX
PA   (SHOK-) SHOKUHIN SANGYO KOS.
XX
DR   WPI; 1991-299435/41.
DR   N-PSDB; AAQ13852.
XX
PT   Neutral protease II gene from Aspergillus - used to produce
PT   recombinant enzyme by expression in Saccharomyces.
XX
PS   Claim; Fig 4; 15pp; Japanese.
XX
CC   The pre-pro neutral protease II is derived from Aspergillus
CC   oryzae and is the precursor to neutral protease II. The neutral
CC   protease II can be expressed in Saccharomyces cerevisae and can be
CC   produced efficiently and in a secreted form.
XX
SQ   Sequence   352 AA;


  Query Match             90.5%;  Score 19;  DB 12;  Length 352;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    240 eagsts 245


RESULT  37
AAU36356
ID   AAU36356 standard; Protein; 363 AA.
XX
AC   AAU36356;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa cellular proliferation protein #346.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS54215.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 11949; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
```

```
SQ   Sequence   363 AA;

  Query Match             90.5%;  Score 19;  DB 22;  Length 363;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    196 eagass 201


RESULT  38
AAY05663
ID   AAY05663 standard; Protein; 366 AA.
XX
AC   AAY05663;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Maize caffeic O-methyltransferase.
XX
KW   Maize; corn; caffeic O-methyltransferase; lignin; transgenic plant.
XX
OS   Zea mays.
XX
PN   WO9910498-A2.
XX
PD   04-MAR-1999.
XX
PF   24-AUG-1998;   98WO-US17519.
XX
PR   12-MAY-1998;   98US-0076851.
PR   27-AUG-1997;   97US-0057082.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bowen BA,  Helentjaris TG,  Wang X;
XX
DR   WPI; 1999-204667/17.
DR   N-PSDB; AAX25202.
XX
PT   Nucleic acids encoding plant lignin biosynthesis enzymes - used to
PT   transform plants to modulate lignin biosynthesis
XX
PS   Claim 9; Page 96-97; 166pp; English.
XX
CC   The present sequence is a caffeic O-methyltransferase of maize,
CC   encoded by a clone (see AAX25202) isolated from a maize line B73 cDNA
CC   library.  The invention provides methods and compositions relating
CC   to altering lignin biosynthesis and/or the lignin composition of
CC   plants.  Isolated nucleic acids (see AAX25196-216) that code for
CC   proteins (see AAY05657-77) involved in lignin biosynthesis are
CC   claimed.  Also claimed are recombinant expression cassettes, host
CC   cells (especially maize or sorghum), and transgenic plants and
CC   seeds.  The claimed nucleic acids can be used to transform a plant
CC   to modulate lignin biosynthesis.  A claimed method involves
CC   transforming a plant cell with a recombinant expression cassette
CC   comprising a lignin biosynthesis polynucleotide operably linked to
CC   a promoter, growing the plant cell under plant growing conditions,
CC   and inducing expression of the polynucleotide for a time sufficient
CC   to modulate (preferably increase) lignin biosynthesis in the plant.
CC   The plant lignins can be used as chemical feedstock.  Plant
CC   material of increased lignin content can be used as a fuel source,
CC   and in the pulp and paper industry.  Decreased lignin content
CC   improves the digestibility of fodder crops.
XX
SQ   Sequence   366 AA;

  Query Match             90.5%;  Score 19;  DB 20;  Length 366;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    115 eagtas 120


RESULT  39
AAG20561
ID   AAG20561 standard; Protein; 377 AA.
XX
AC   AAG20561;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 22802.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
```

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 377;
Best Local Similarity 66.7%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 75 eagsss 80

RESULT 40

AAG41800
ID AAG41800 standard; Protein; 377 AA.
XX
AC AAG41800;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52052.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

```
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             90.5%;  Score 19;  DB 21;  Length 377;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      75 eagsss 80


RESULT  41
AAG24042
ID   AAG24042 standard; Protein; 386 AA.
XX
AC   AAG24042;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 27567.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
```

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 386;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 eaqxxs 6
           |||  |
Db      46 eagaas 51
```

RESULT 42
AAP70581
ID AAP70581 standard; Protein; 390 AA.
XX
AC AAP70581;
XX
DT 03-MAY-1991 (first entry)
XX
DE Protease biosynthetic protein.
XX
KW Saccharomycosis; yeast.
XX
PN JP62104578-A.
XX
PD 15-MAY-1987.

```
XX
PF   31-OCT-1985;   85JP-0244893.
XX
PR   31-OCT-1985;   85JP-0244893.
XX
PA   (FUKU/) FUKUI S.
XX
DR   WPI; 1987-173695/25.
DR   N-PSDB; AAN70927.
XX
PT   Protease prodn. - by culturing microorganism transformed with
PT   vector derived from saccharomycosis.
XX
PS   Disclosure; Fig 1; 9pp; Japanese.
XX
CC   Product is a biosynthetic component involved in the synthesis of
CC   protease. The protein may be produced from a transformed S.cerevisiae
CC   expression system for the large scale production of protease.
XX
SQ   Sequence   390 AA;


  Query Match             90.5%;  Score 19;  DB 8;  Length 390;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    266 eagsss 271


RESULT  43
AAY35147
ID   AAY35147 standard; Protein; 393 AA.
XX
AC   AAY35147;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae transmembrane protein sequence.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1016; Disclosure; 1912pp; English.
XX
CC   AAY34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   393 AA;


  Query Match             90.5%;  Score 19;  DB 20;  Length 393;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    101 eagsss 106


RESULT  44
AAY28643
ID   AAY28643 standard; Protein; 422 AA.
XX
AC   AAY28643;
XX
DT   03-NOV-1999  (first entry)
XX
DE   Human serine protease inhibitor from cDNA clone HETDK50.
XX
KW   Human serine protease inhibitor from cDNA clone HETDK50; fusion protein;
KW   serpin; serine protease; human pre-alpha-1-antitrypsin precursor;
KW   extracellular matrix degradation; multiple sclerosis; cancer; arthritis;
KW   inflammation; immune system disorder; neurodegenerative disorder;
KW   Kallmann's syndrome; Down's syndrome; Alzheimer's; secreted protein;
KW   galactorrhea; hypogonadism; somatostatin; protein purification.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /label= Signal_peptide
FT   Protein         20..422
FT                   /label= Mature_serine_protease_inhibitor
XX
PN   WO9940183-A1.
XX
PD   12-AUG-1999.
XX
PF   04-FEB-1999;   99WO-US02292.
XX
PR   06-FEB-1998;   98US-0073961.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Ruben SM;
XX
DR   WPI; 1999-508502/42.
DR   N-PSDB; AAX80907.
XX
PT   New isolated human serine protease and serpin polypeptides, used to
PT   develop products for treating e.g. immune disorders, cancers,
PT   inflammation, transplant rejection or infections, or as food
PT   additives
XX
PS   Claim 11; Pages 83-85; 99pp; English.
XX
CC   The present sequence is a serine protease inhibitor (serpin) from cDNA
CC   clone HETDK50 which is obtained from human endometrial tumour tissue
CC   cDNA library. The protein shows a high degree of sequence similarity to
CC   human pre-alpha-1-antitrypsin precursor. The serpin and its
CC   coding sequence are used in the diagnosis and treatment of disorders
CC   related to abnormal level of the protein or mutation in the nucleotide
CC   sequence. The serpin can be used for treating disorders characterised by
CC   degradation of extracellular matrix, e.g. cancer, arthritis, multiple
CC   sclerosis and immune system disorders, for treating wasting associated
```

```
CC   with excessive protease production during inflammation or
CC   neurodegenerative disorders e.g. Kallmann's and Down's syndromes,
CC   Alzheimer's and Huntington's diseases. It may also be used to reduce
CC   excess levels of prolactin in the treatment of galactorrhea and
CC   hypogonadism, and decrease the amount of free circulating somatostatin to
CC   prevent somatostatin's inhibitory effect on the release of growth
CC   hormone. The fusion of this protein to His-tag, HA-tag, IgG domains,
CC   etc. facilitates protein purification and fusion to IgG-1, IgG-3 and
CC   albumin increases the half life time in vivo.
XX
SQ   Sequence   422 AA;


  Query Match             90.5%;  Score 19;  DB 20;  Length 422;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    370 eagaas 375


RESULT  45
AAB74691
ID   AAB74691 standard; Protein; 422 AA.
XX
AC   AAB74691;
XX
DT   12-JUN-2001  (first entry).
XX
DE   Human protease and protease inhibitor PPIM-24.
XX
KW   Human; protease; protease inhibitor; protease and protease inhibitor;
KW   PPIM; identification; diagnosis; anti-human immunodeficiency virus; HIV;
KW   antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW   antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW   cytostatic; antibacterial; fungicide; protozoacide; antiarteriosclerotic;
KW   antiatherosclerotic; antipsoriatic; virucide; hepatotropic; gene therapy;
KW   autoimmune disorder; inflammatory disorder; AIDS; DiGeorge's syndrome;
KW   severe combined immunodeficiency disease; SCID; Chediak-Higashi syndrome;
KW   Cushing's disease; Addison's disease; autoimmune thyroiditis; gout;
KW   Crohn's disease; diabetes mellitus; Good pasture's syndrome; infection;
KW   Grave's diseases; Hashimoto's thyroiditis; Sjogren's syndrome; cancer;
KW   Werner's syndrome; cell proliferative disorder; arteriosclerosis;
KW   atherosclerosis; cirrhosis; hepatitis; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200110903-A2.
XX
PD   15-FEB-2001.
XX
PF   09-AUG-2000; 2000WO-US21878.
XX
PR   09-AUG-1999;   99US-0147986.
PR   21-OCT-1999;   99US-0160807.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Lal P,  Tang YT,  Bandman O,  Baughn MR,  Azimzai Y,  Lu DAM;
PI   Yang J;
XX
DR   WPI; 2001-202760/20.
DR   N-PSDB; AAF81737.
XX
PT   New protease (inhibitors) useful for diagnosis and treatment of
PT   autoimmune/inflammatory disorders such as acquired immunodeficiency
PT   syndrome, Cushing's disease, Addison's disease and cell proliferative
PT   disorders such as cancer  -
XX
PS   Claim 1; Page 112-113; 134pp; English.
XX
CC   AAF81714 to AAF81740 encode the human proteases and protease inhibitors
CC   (PPIMs) given in AAB74668 to AAB74694. The PPIMs can have activities such
CC   as: anti-human immunodeficiency virus (HIV); antidiabetic; antithyroid;
CC   immunostimulant; immunomodulator; antiinflammatory; immunosuppressive;
CC   nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;
CC   fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;
CC   virucide; antipsoriatic; and hepatotropic. PPIM polynucleotide and
CC   protein sequences can be used in the diagnosis, treatment and prevention
CC   of autoimmune/inflammatory disorders such as AIDS, DiGeorge's syndrome,
CC   severe combined immunodeficiency disease (SCID), Chediak-Higashi
CC   syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,
CC   Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,
CC   Grave's diseases, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's
CC   syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic
CC   infections and cell proliferative disorder such as arteriosclerosis,
CC   atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPIM
CC   polynucleotide sequences can be used in somatic or germline gene therapy
CC   and in diagnosis of diseases. They can also be used in generating
CC   hybridisation probes useful in mapping the naturally occurring genomic
CC   sequences and in molecular biology techniques.
XX
SQ   Sequence   422 AA;


  Query Match             90.5%;  Score 19;  DB 22;  Length 422;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    370 eagaas 375


RESULT  46
AAG41388
ID   AAG41388 standard; Protein; 427 AA.
XX
AC   AAG41388;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 51488.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
```

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 427;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 eagxxs 6
          |||  |
Db      6 eagsss 11
```

RESULT 47
AAG41387
ID AAG41387 standard; Protein; 431 AA.
XX
AC AAG41387;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51487.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 19; DB 21; Length 431;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 eagxxs 6
           |||  |
Db      10 eagsss 15
```

RESULT 48
AAG24041
ID AAG24041 standard; Protein; 435 AA.
XX
AC AAG24041;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27566.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```
PR   29-OCT-1999;   99US-0162142.

  Query Match            90.5%;  Score 19;  DB 21;  Length 435;
  Best Local Similarity  66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      95 eagaas 100


RESULT  49
AAG24040
ID   AAG24040 standard; Protein; 440 AA.
XX
AC   AAG24040;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 27565.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
```

```
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           90.5%;  Score 19;  DB 21;  Length 440;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
```

```
          |||  |
Db    100 eagaas 105


RESULT  50
AAR31955
ID   AAR31955 standard; Protein; 441 AA.
XX
AC   AAR31955;
XX
DT   06-JUN-1993  (first entry)
XX
DE   Sequence encoded by glycoprotein G gene.
XX
KW   IBR glycoprotein E gene; unique short 2 gene.
XX
OS   Infectious bovine rhinotracheitis.
XX
PN   WO9302104-A.
XX
PD   04-FEB-1993.
XX
PF   20-JUL-1992;   92WO-US06034.
XX
PR   18-JUL-1991;   91US-0732584.
XX
PA   (SYTR ) SYNTRO CORP.
XX
PI   Cochran MD,  Macdonald RD;
XX
DR   WPI; 1993-058725/07.
DR   N-PSDB; AAQ36768.
XX
PT   Recombinant infectious bovine rhinotracheitis virus - provides
PT   isolated DNA encoding gpE glyco:protein, gpG glyco:protein and
PT   unique short 2 genes of the virus
XX
PS   Example; Fig 8; 240pp; English.
XX
CC   The sequence of approximately 1400 base pairs of the HindIII K
CC   fragment, starting approximately 2800 base pairs downstream of the
CC   HindIII K/HindIII O junction, are shown.  The glycoprotein G (gpG)
CC   gene is transcribed away from the HindIII K/HindIII O junction.
XX
SQ   Sequence   441 AA;


  Query Match           90.5%;  Score 19;  DB 14;  Length 441;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    265 eagsas 270


Search completed: August 30, 2002, 15:05:43
Job time: 6419 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 30, 2002, 12:42:09 ; Search time 21.5 Seconds
(without alignments)
6.816 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	90.5	47	4	US-08-936-165A-345	Sequence 345, App
2	19	90.5	159	4	US-08-856-253-2	Sequence 2, Appli
3	19	90.5	211	4	US-08-856-253-4	Sequence 4, Appli
4	19	90.5	332	2	US-08-405-175A-5	Sequence 5, Appli
5	19	90.5	441	1	US-08-191-866D-21	Sequence 21, Appl
6	19	90.5	441	2	US-08-185-949B-21	Sequence 21, Appl
7	19	90.5	512	4	US-08-856-253-6	Sequence 6, Appli
8	19	90.5	908	5	PCT-US95-03747-3	Sequence 3, Appli
9	19	90.5	1088	4	US-09-130-242-2	Sequence 2, Appli
10	19	90.5	1183	2	US-08-447-031A-2	Sequence 2, Appli
11	18	85.7	9	4	US-08-485-355B-31	Sequence 31, Appl
12	18	85.7	59	2	US-08-596-387B-109	Sequence 109, App
13	18	85.7	59	4	US-09-067-615-109	Sequence 109, App
14	18	85.7	59	5	PCT-US95-09816A-109	Sequence 109, App
15	18	85.7	61	4	US-08-446-137B-12	Sequence 12, Appl
16	18	85.7	66	4	US-08-446-137B-10	Sequence 10, Appl
17	18	85.7	91	2	US-08-598-873-19	Sequence 19, Appl
18	18	85.7	91	4	US-08-605-430-19	Sequence 19, Appl
19	18	85.7	123	1	US-07-869-912-2	Sequence 2, Appli
20	18	85.7	175	4	US-08-446-137B-9	Sequence 9, Appli
21	18	85.7	211	2	US-08-708-958-2	Sequence 2, Appli
22	18	85.7	428	4	US-09-423-340-2	Sequence 2, Appli
23	18	85.7	433	2	US-08-749-902-1	Sequence 1, Appli
24	18	85.7	481	3	US-09-330-095-1	Sequence 1, Appli
25	18	85.7	500	4	US-08-960-190A-25	Sequence 25, Appl
26	18	85.7	503	1	US-07-612-673-2	Sequence 2, Appli
27	18	85.7	503	4	US-08-845-258-52	Sequence 52, Appl
28	18	85.7	503	4	US-08-990-571-52	Sequence 52, Appl
29	18	85.7	521	1	US-07-796-361A-11	Sequence 11, Appl
30	18	85.7	521	1	US-08-539-666-2	Sequence 2, Appli
31	18	85.7	525	1	US-08-348-891A-2	Sequence 2, Appli
32	18	85.7	525	2	US-08-905-817-2	Sequence 2, Appli
33	18	85.7	530	1	US-08-513-841-1	Sequence 1, Appli
34	18	85.7	530	2	US-08-696-834-1	Sequence 1, Appli
35	18	85.7	530	2	US-08-942-673-1	Sequence 1, Appli
36	18	85.7	530	4	US-09-118-317-1	Sequence 1, Appli
37	18	85.7	560	2	US-08-808-931-18	Sequence 18, Appl
38	18	85.7	560	3	US-08-808-323-18	Sequence 18, Appl
39	18	85.7	560	3	US-09-050-603A-18	Sequence 18, Appl
40	18	85.7	560	3	US-09-102-420B-18	Sequence 18, Appl
41	18	85.7	560	4	US-09-497-698-18	Sequence 18, Appl
42	18	85.7	566	1	US-08-419-078-2	Sequence 2, Appli
43	18	85.7	566	1	US-08-726-883-2	Sequence 2, Appli
44	18	85.7	730	2	US-08-696-944-2	Sequence 2, Appli
45	18	85.7	731	1	US-08-731-716-2	Sequence 2, Appli
46	18	85.7	823	4	US-09-651-656-19	Sequence 19, Appl
47	18	85.7	968	1	US-08-434-730-14	Sequence 14, Appl
48	18	85.7	976	3	US-08-560-005-2	Sequence 2, Appli
49	18	85.7	976	3	US-09-195-868-14	Sequence 14, Appl
50	18	85.7	976	4	US-09-418-540-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-936-165A-345
; Sequence 345, Application US/08936165A
; Patent No. 6348582
;  GENERAL INFORMATION:
;    APPLICANT:  Black, Michael
;    APPLICANT:  Burnham, Martin
;    APPLICANT:  Hodgson, John
;    APPLICANT:  Knowles, David
;    APPLICANT:  Lonetto, Michael
;    APPLICANT:  Nicholas, Richard
;    APPLICANT:  Pratt, Julie
;    APPLICANT:  Reichard, Richard
;    APPLICANT:  Rosenberg,.Martin
;    APPLICANT:  Ward, Judith
;    TITLE OF INVENTION:  No. 6348582el Prokaryotic Polynucleotides,
;    TITLE OF INVENTION:  Polypeptides and Their Uses
;    NUMBER OF SEQUENCES:  534
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/936,165A
;      FILING DATE:  24-SEP-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/027,032
;      FILING DATE:  24-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gimmi, Edward R
;      REGISTRATION NUMBER:  38,891
;      REFERENCE/DOCKET NUMBER:  P50549
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-4478
;      TELEFAX:  610-270-5090
;      TELEX:
```

```
;  INFORMATION FOR SEQ ID NO:  345:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-08-936-165A-345


  Query Match             90.5%;  Score 19;  DB 4;  Length 47;
  Best Local Similarity   66.7%;  Pred. No. 93;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     31 EAGATS 36


RESULT   2
US-08-856-253-2
; Sequence 2, Application US/08856253
; Patent No. 6288214
;  GENERAL INFORMATION:
;    APPLICANT:  Hook, Magnus
;    APPLICANT:  Patti, Joseph M.
;    APPLICANT:  House-Pompeo, Karen
;    APPLICANT:  Sthanam, Narayana
;    APPLICANT:  Symersky, Jindrich
;    TITLE OF INVENTION:  COLLAGEN BINDING PROTEIN COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/856,253
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/017,678
;      FILING DATE:  16-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  TAMK:193
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  159 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-856-253-2


  Query Match             90.5%;  Score 19;  DB 4;  Length 159;
  Best Local Similarity   66.7%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     28 EAGTSS 33


RESULT   3
US-08-856-253-4
; Sequence 4, Application US/08856253
; Patent No. 6288214
;  GENERAL INFORMATION:
;    APPLICANT:  Hook, Magnus
;    APPLICANT:  Patti, Joseph M.
;    APPLICANT:  House-Pompeo, Karen
;    APPLICANT:  Sthanam, Narayana
;    APPLICANT:  Symersky, Jindrich
;    TITLE OF INVENTION:  COLLAGEN BINDING PROTEIN COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/856,253
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/017,678
;      FILING DATE:  16-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  TAMK:193
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  211 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-856-253-4


  Query Match             90.5%;  Score 19;  DB 4;  Length 211;
  Best Local Similarity   66.7%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     34 EAGTSS 39


RESULT   4
US-08-405-175A-5
; Sequence 5, Application US/08405175A
; Patent No. 5885772
;  GENERAL INFORMATION:
;    APPLICANT:  Aderem, Alan A.
;    APPLICANT:  Chen, Jianmin
;    APPLICANT:  Chang, Sandy
;    TITLE OF INVENTION:  METHOD FOR THE DETECTION OF ANENCEPHALY
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
```

```
;       CITY:  Hackensack
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/405,175A
;       FILING DATE:  16-MAR-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-121A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  332 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;       DESCRIPTION:  predicted primary structure of human MARCKS
;     HYPOTHETICAL:  NO
US-08-405-175A-5


  Query Match             90.5%;  Score 19;  DB 2;  Length 332;
  Best Local Similarity   66.7%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    206 EAGAAS 211


RESULT   5
US-08-191-866D-21
; Sequence 21, Application US/08191866D
; Patent No. 5783195
;  GENERAL INFORMATION:
;     APPLICANT:  Cochran, Mark D
;     APPLICANT:  Macdonald, Richard D.
;     TITLE OF INVENTION:  Recombinant Infectious Bovine
;     TITLE OF INVENTION:  Rhinotracheitis Virus S-IBR-052 And Uses Thereof
;     NUMBER OF SEQUENCES:  99
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  John P. White
;       STREET:  1185 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/191,866D
;       FILING DATE:  4 February 1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White, John P.
;       REGISTRATION NUMBER:  28,678
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 278-0400
;       TELEFAX:  (212) 391-0525
;       TELEX:  422523
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  441 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-191-866D-21


  Query Match             90.5%;  Score 19;  DB 1;  Length 441;
  Best Local Similarity   66.7%;  Pred. No. 8.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    265 EAGSAS 270


RESULT   6
US-08-185-949B-21
; Sequence 21, Application US/08185949B
; Patent No. 5874279
;  GENERAL INFORMATION:
;     APPLICANT:  Mark D. Cochran
;     APPLICANT:  Richard D. Macdonald
;     TITLE OF INVENTION:  Recombinant Infectious Bovine
;     TITLE OF INVENTION:  Rhinotracheitis Virus
;     NUMBER OF SEQUENCES:  104
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  John P. White
;       STREET:  1185 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM 330 466 DX2
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/185,949B
;       FILING DATE:  03-NOV-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White, John P.
;       REGISTRATION NUMBER:  ,678
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 278-0400
;       TELEFAX:  (212) 278-0525
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  441 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-185-949B-21


  Query Match             90.5%;  Score 19;  DB 2;  Length 441;
  Best Local Similarity   66.7%;  Pred. No. 8.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    265 EAGSAS 270


RESULT   7
US-08-856-253-6
; Sequence 6, Application US/08856253
```

```
; Patent No. 6288214
;  GENERAL INFORMATION:
;    APPLICANT:  Hook, Magnus
;    APPLICANT:  Patti, Joseph M.
;    APPLICANT:  House-Pompeo, Karen
;    APPLICANT:  Sthanam, Narayana
;    APPLICANT:  Symersky, Jindrich
;    TITLE OF INVENTION:  COLLAGEN BINDING PROTEIN COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/856,253
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/017,678
;      FILING DATE:  16-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  TAMK:193
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  512 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-856-253-6


  Query Match             90.5%;  Score 19;  DB 4;  Length 512;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     149 EAGTSS 154


RESULT   8
PCT-US95-03747-3
; Sequence 3, Application PC/TUS9503747
;  GENERAL INFORMATION:
;    APPLICANT:  LA JOLLA CANCER RESEARCH FOUNDATION
;    TITLE OF INVENTION:  Brevican, A Glial Cell Proteoglycan
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03747
;      FILING DATE:  27-MAR-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Imbra, Richard J.
;      REGISTRATION NUMBER:  37,643
;      REFERENCE/DOCKET NUMBER:  FP-LJ 1453
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  908 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
PCT-US95-03747-3


  Query Match             90.5%;  Score 19;  DB 5;  Length 908;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     583 EAGSSS 588


RESULT   9
US-09-130-242-2
; Sequence 2, Application US/09130242B
; Patent No. 6194558
; GENERAL INFORMATION:
;  APPLICANT: Gianturco, S.H.
;  APPLICANT: Bradley, W.A.
;  TITLE OF INVENTION: DNA Encoding Human Monocyte-Macrophage Aoplipoprotein
;  TITLE OF INVENTION: B Receptor Gene and Protein
;  FILE REFERENCE: D5880
;  CURRENT APPLICATION NUMBER: US/09/130,242B
;  CURRENT FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: MS WORD, Macintosh OS 8.5
; SEQ ID NO 2
;   LENGTH: 1088
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-130-242-2


  Query Match             90.5%;  Score 19;  DB 4;  Length 1088;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     327 EAGTAS 332


RESULT  10
US-08-447-031A-2
; Sequence 2, Application US/08447031A
; Patent No. 5851794
;  GENERAL INFORMATION:
;    APPLICANT:  GUSS, Bengt
;    APPLICANT:  HOOK, Magnus
;    APPLICANT:  JONSSON, Hans
;    APPLICANT:  LINDBERG, Martin
;    APPLICANT:  PATTI, Joseph
;    APPLICANT:  SIGNAS, Christer
;    APPLICANT:  SWITALSKI, Lech
;    TITLE OF INVENTION:  A COLLAGEN BINDING PROTEIN AS WELL AS
;    TITLE OF INVENTION:  ITS PREPARATION
;    NUMBER OF SEQUENCES:  8
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,031A
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/861,804
;      FILING DATE:  21-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/SE91/00707
;      FILING DATE:  22-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SE 9003374-7
;      FILING DATE:  22-OCT-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McGowan, Malcolm K.
;      REGISTRATION NUMBER:  39,300
;      REFERENCE/DOCKET NUMBER:  012889-006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1183 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-447-031A-2


  Query Match             90.5%;  Score 19;  DB 2;  Length 1183;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     166 EAGTSS 171


RESULT  11
US-08-485-355B-31
; Sequence 31, Application US/08485355B
; Patent No. 6177075
;   GENERAL INFORMATION:
;        APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
;        TITLE OF INVENTION: Insect Viruses and Their Uses in
;                            Protecting Plants
;        NUMBER OF SEQUENCES: 57
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
;             STREET: Four Embarcadero Center, Suite 3400
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States
;             ZIP: 94111-4187
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/485,355B
;             FILING DATE: 07-Jun-1995
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/440,522
;             FILING DATE: 12-MAY-1995
;             APPLICATION NUMBER: US 08/089,372
;             FILING DATE: 08-JUL-1993
;             APPLICATION NUMBER: AU PL4081/92
;             FILING DATE: 14-AUG-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Trecartin, Richard F.
;             REGISTRATION NUMBER: 31,801
;             REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 781-1989
;             TELEFAX: (415) 398-3249
;             TELEX: 910 277299
;   INFORMATION FOR SEQ ID NO: 31:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-08-485-355B-31


  Query Match             85.7%;  Score 18;  DB 4;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db       3 EAGVAS 8


RESULT  12
US-08-596-387B-109
; Sequence 109, Application US/08596387B
; Patent No. 5869270
;  GENERAL INFORMATION:
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Jiao, Jin-An
;    APPLICANT:  Burkhardt, Martin
;    APPLICANT:  Wong, Hing
;    TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dade International, Inc.
;      STREET:  1717 Deerfield Road
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/596,387B
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09816
;      FILING DATE:  31-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/382,454
;      FILING DATE:  01-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/283,302
;      FILING DATE:  29-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Pearson, Louise S.
;      REGISTRATION NUMBER:  32,369
;      REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708) 267-5300
;      TELEFAX:  (708) 267-5376
;  INFORMATION FOR SEQ ID NO:  109:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-08-596-387B-109


  Query Match             85.7%;  Score 18;  DB 2;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     41 EAGRAS 46


RESULT  13
US-09-067-615-109
; Sequence 109, Application US/09067615
; Patent No. 6309645
;  GENERAL INFORMATION:
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Jiao, Jin-An
;    APPLICANT:  Burkhardt, Martin
;    APPLICANT:  Wong, Hing
;    TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dade International, Inc.
;      STREET:  1717 Deerfield Road
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/067,615
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/596,387
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/382,454
;      FILING DATE:  01-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/283,302
;      FILING DATE:  29-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pearson, Louise S.
;      REGISTRATION NUMBER:  32,369
;      REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708) 267-5300
;      TELEFAX:  (708) 267-5376
;  INFORMATION FOR SEQ ID NO:  109:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-09-067-615-109


  Query Match             85.7%;  Score 18;  DB 4;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     41 EAGRAS 46


RESULT  14
PCT-US95-09816A-109
; Sequence 109, Application PC/TUS9509816A
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, Hing C.
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Widanz, Jon A.
;    APPLICANT:  Grammer, Susan
;    APPLICANT:  Edwards, Ana C.
;    APPLICANT:  Chavaillaz, Pierre-Andre
;    APPLICANT:  Jiao, Jin-An
;    TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;    NUMBER OF SEQUENCES:  123
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dade International, Inc.
;      STREET:  1717 Deerfield Road
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09816A
;      FILING DATE:  31-JUL-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/382,454
;      FILING DATE:  01-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/283,302
;      FILING DATE:  29-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pearson, Louise S.
;      REGISTRATION NUMBER:  32,369
;      REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708) 267-5300
;      TELEFAX:  (708) 267-5376
;  INFORMATION FOR SEQ ID NO:  109:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
PCT-US95-09816A-109


  Query Match             85.7%;  Score 18;  DB 5;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     41 EAGRAS 46


RESULT  15
```

```
US-08-446-137B-12
; Sequence 12, Application US/08446137B
; Patent No. 6162903
;  GENERAL INFORMATION:
;    APPLICANT:  Trowern, Angus R.
;    APPLICANT:  Atkinson, Anthony
;    APPLICANT:  Murphy, Jonathan P.
;    APPLICANT:  Laurence, Oliver S.
;    APPLICANT:  Duggleby, Clive J.
;    TITLE OF INVENTION:  IMMUNOGLOBULIN BINDING PROTEINS DERIVED
;    TITLE OF INVENTION:  FROM L PROTEIN AND THEIR USES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,137B
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  100084.406
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-446-137B-12


  Query Match             85.7%;  Score 18;  DB 4;  Length 61;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      29 EAGITS 34


RESULT  16
US-08-446-137B-10
; Sequence 10, Application US/08446137B
; Patent No. 6162903
;  GENERAL INFORMATION:
;    APPLICANT:  Trowern, Angus R.
;    APPLICANT:  Atkinson, Anthony
;    APPLICANT:  Murphy, Jonathan P.
;    APPLICANT:  Laurence, Oliver S.
;    APPLICANT:  Duggleby, Clive J.
;    TITLE OF INVENTION:  IMMUNOGLOBULIN BINDING PROTEINS DERIVED
;    TITLE OF INVENTION:  FROM L PROTEIN AND THEIR USES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,137B
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  100084.406
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  66 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-446-137B-10


  Query Match             85.7%;  Score 18;  DB 4;  Length 66;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      33 EAGISS 38


RESULT  17
US-08-598-873-19
; Sequence 19, Application US/08598873
; Patent No. 5928884
;  GENERAL INFORMATION:
;    APPLICANT:  Croce, Carlo M.
;    APPLICANT:  Huebner, Kay
;    TITLE OF INVENTION:  FHIT PROTEINS AND NUCLEIC ACIDS AND
;    TITLE OF INVENTION:  METHODS BASED THEREON
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/598,873
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Friebel, Thomas E.
;      REGISTRATION NUMBER:  29,258
;      REFERENCE/DOCKET NUMBER:  8666-004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  91 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
```

```
;    MOLECULE TYPE:  peptide
US-08-598-873-19


  Query Match             85.7%;  Score 18;  DB 2;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 3.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     49 EAGKSS 54


RESULT  18
US-08-605-430-19
; Sequence 19, Application US/08605430
; Patent No. 6242212
;  GENERAL INFORMATION:
;    APPLICANT:  Croce, Carlo M.
;    APPLICANT:  Huebner, Kay
;    TITLE OF INVENTION:  FHIT PROTEINS AND NUCLEIC ACIDS AND
;    TITLE OF INVENTION:  METHODS BASED THEREON
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,430
;      FILING DATE:  22-FEB-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Friebel, Thomas E.
;      REGISTRATION NUMBER:  29,258
;      REFERENCE/DOCKET NUMBER:  8666-005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  91 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-605-430-19


  Query Match             85.7%;  Score 18;  DB 4;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 3.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     49 EAGKSS 54


RESULT  19
US-07-869-912-2
; Sequence 2, Application US/07869912
; Patent No. 5316922
;  GENERAL INFORMATION:
;    APPLICANT:  Court, Don
;    APPLICANT:  Brown, Stanley
;    TITLE OF INVENTION:  A Method for Identifying and
;    TITLE OF INVENTION:  Expressing Proteins that Recognize and Adhere to Specific
;    TITLE OF INVENTION:  Probes
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenneth A. Weber
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/869,912
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Kenneth A.
;      REGISTRATION NUMBER:  32,334
;      REFERENCE/DOCKET NUMBER:  15280-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-543-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  123 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-869-912-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 123;
  Best Local Similarity   66.7%;  Pred. No. 4.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     21 EAGGSS 26


RESULT  20
US-08-446-137B-9
; Sequence 9, Application US/08446137B
; Patent No. 6162903
;  GENERAL INFORMATION:
;    APPLICANT:  Trowern, Angus R.
;    APPLICANT:  Atkinson, Anthony
;    APPLICANT:  Murphy, Jonathan P.
;    APPLICANT:  Laurence, Oliver S.
;    APPLICANT:  Duggleby, Clive J.
;    TITLE OF INVENTION:  IMMUNOGLOBULIN BINDING PROTEINS DERIVED
;    TITLE OF INVENTION:  FROM L PROTEIN AND THEIR USES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,137B
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  514
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  100084.406
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  175 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-446-137B-9


  Query Match             85.7%;  Score 18;  DB 4;  Length 175;
  Best Local Similarity   66.7%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      29 EAGITS 34


RESULT  21
US-08-708-958-2
; Sequence 2, Application US/08708958
; Patent No. 5948952
;  GENERAL INFORMATION:
;    APPLICANT:  SANDS, Arthur T.
;    APPLICANT:  BRADLEY, Allan
;    APPLICANT:  ABUIN, Alejandro
;    TITLE OF INVENTION:  XERODERMA PIGMENTOSUM-DEFICIENT
;    TITLE OF INVENTION:  MOUSE
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 PENNSYLVANIA AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/708,958
;      FILING DATE:  __ SEP-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIT, Gordon
;      REGISTRATION NUMBER:  30,764
;      REFERENCE/DOCKET NUMBER:  A-6641
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  211 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-708-958-2


  Query Match             85.7%;  Score 18;  DB 2;  Length 211;
  Best Local Similarity   66.7%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      25 EAGSGS 30


RESULT  22
US-09-423-340-2
; Sequence 2, Application US/09423340
; Patent No. 6225454
; GENERAL INFORMATION:
;  APPLICANT: MIYAGI, Taeko
;  APPLICANT: WADA, Tadashi
;  APPLICANT: YOSHIKAWA, Yuko
;  TITLE OF INVENTION: SIALIDASE LOCALIZED IN PLASMA MEMBRANE AND
;  TITLE OF INVENTION: DNA CODING FOR THE SAME
;  FILE REFERENCE: OP699
;  CURRENT APPLICATION NUMBER: US/09/423,340
;  CURRENT FILING DATE: 1999-11-22
;  EARLIER APPLICATION NUMBER: JP 9-132174
;  EARLIER FILING DATE: 1997-05-22
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Bos primigenius
US-09-423-340-2


  Query Match             85.7%;  Score 18;  DB 4;  Length 428;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     321 EAGTLS 326


RESULT  23
US-08-749-902-1
; Sequence 1, Application US/08749902
; Patent No. 5985635
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  NOVEL HUMAN SERINE/THREONINE
;    TITLE OF INVENTION:  PROTEIN KINASES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,902
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0150 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
```

```
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  433 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  Consensus
US-08-749-902-1


  Query Match             85.7%;  Score 18;  DB 2;  Length 433;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    396 EAGAVS 401


RESULT  24
US-09-330-095-1
; Sequence 1, Application US/09330095
; Patent No. 6127161
; GENERAL INFORMATION:
;  APPLICANT: Kikkoman Corporation
;  TITLE OF INVENTION: Leucine Aminopeptidase Gene, Recombinant DNA, and
;  TITLE OF INVENTION: Process for Producing Leucine Aminopeptidase
;  FILE REFERENCE: PH-622
;  CURRENT APPLICATION NUMBER: US/09/330,095
;  CURRENT FILING DATE: 1999-06-11
;  EARLIER APPLICATION NUMBER: JP-164611/1998
;  EARLIER FILING DATE: 1998-06-12
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Aspergillus sojae
US-09-330-095-1


  Query Match             85.7%;  Score 18;  DB 3;  Length 481;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    193 EAGSVS 198


RESULT  25
US-08-960-190A-25
; Sequence 25, Application US/08960190A
; Patent No. 6232445
;  GENERAL INFORMATION:
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Acevedo, Jorge
;    APPLICANT:  Burkhardt, Martin
;    APPLICANT:  Jiao, Jin-an
;    APPLICANT:  Wong, Hing C.
;    TITLE OF INVENTION:  SOLUBLE MHC COMPLEXES AND
;    TITLE OF INVENTION:  METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dike, Bronstein, Roberts & Cushman, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  usa
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/960,190A
;      FILING DATE:  29-OCT-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corless, Peter F
;      REGISTRATION NUMBER:  33,860
;      REFERENCE/DOCKET NUMBER:  48002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  500 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-960-190A-25


  Query Match             85.7%;  Score 18;  DB 4;  Length 500;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     41 EAGRAS 46


RESULT  26
US-07-612-673-2
; Sequence 2, Application US/07612673
; Patent No. 5260208
;  GENERAL INFORMATION:
;    APPLICANT:  Petre, Dominique
;    APPLICANT:  Cerbelaud, Edith
;    APPLICANT:  Mayaux, Jean-Francois
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NOVEL POLYPEPTIDES, THE DNA SEQUENCES
;    TITLE OF INVENTION:  ALLOWING THEIR EXPRESSION, METHOD OF PREPARATION, AND
;    TITLE OF INVENTION:  UTILIZATION
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/612,673
;      FILING DATE:  19901114
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Potter, Jane E.R.
;      REGISTRATION NUMBER:  33,332
;      REFERENCE/DOCKET NUMBER:  03715.0010
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (202) 408-4000
;      TELEFAX:  (202) 408-4400
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  503 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-612-673-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 503;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |.
Db     149 EAGGSS 154


RESULT  27
US-08-845-258-52
; Sequence 52, Application US/08845258
; Patent No. 6183976
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Sleath, Paul R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;    TITLE OF INVENTION:  AND TREATMENT OF B. MICROTI INFECTION
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED AND BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,258
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.426C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206)682-6031
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  503 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Babesia Microti
US-08-845-258-52


  Query Match             85.7%;  Score 18;  DB 4;  Length 503;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     417 EAGGTS 422


RESULT  28
US-08-990-571-52
; Sequence 52, Application US/08990571
; Patent No. 6214971
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G. et al.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B.
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED AND BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/990,571
;      FILING DATE:  11-DEC-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206)682-6031
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  503 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Babesia Microti
US-08-990-571-52


  Query Match             85.7%;  Score 18;  DB 4;  Length 503;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     417 EAGGTS 422


RESULT  29
US-07-796-361A-11
; Sequence 11, Application US/07796361A
; Patent No. 5258292
;  GENERAL INFORMATION:
;    APPLICANT:  YEH, Patrice
;    APPLICANT:  MAYAUX, Jean-Francois
;    APPLICANT:  CERBELAUD, Edith
;    APPLICANT:  PETRE, Dominique
;    TITLE OF INVENTION:  ENZYMIC PROCESS FOR THE SYNTHESIS OF
;    TITLE OF INVENTION:  AMMONIUM ADIPATE
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker and Mathis
;      STREET:  The George Mason Building, Washington &
;      STREET:  Prince Streets
;      CITY:  Alexandria
;      STATE:  Virginia
```

```
;       COUNTRY:  United States
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/796,361A
;       FILING DATE:  19911122
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 90-14 853
;       FILING DATE:  28-NOV-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CRANE-FEURY, SHARON E.
;       REGISTRATION NUMBER:  P36,113
;       REFERENCE/DOCKET NUMBER:  003025-010
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 836-6620
;       TELEFAX:  (703) 836-2021
;       TELEX:  440580
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  521 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-796-361A-11


  Query Match             85.7%;  Score 18;  DB 1;  Length 521;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    167 EAGGSS 172


RESULT  30
US-08-539-666-2
; Sequence 2, Application US/08539666
; Patent No. 5766918
;  GENERAL INFORMATION:
;     APPLICANT:  Petre, Dominique
;     APPLICANT:  Cerbelaud, Edith
;     APPLICANT:  Mayaux, Jean-Francois
;     APPLICANT:  Yeh, Patrice
;     TITLE OF INVENTION:  No. 5766918el Polypeptides, The DNA Sequences
;     TITLE OF INVENTION:  Allowing their Expression, Method of Preparation, and
;     TITLE OF INVENTION:  Utilization
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;       ADDRESSEE:  Dunner
;       STREET:  1300 I Street, N.W.
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3315
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/539,666
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/097,009
;       FILING DATE:  27-JUL-1993
;       APPLICATION NUMBER:  US 07/612,673
;       FILING DATE:  14-NOV-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 8916332
;       FILING DATE:  11-DEC-1989
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Potter, Jane E.R.
;       REGISTRATION NUMBER:  33,332
;       REFERENCE/DOCKET NUMBER:  03715.0010-01000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-408-4000
;       TELEFAX:  202-408-4400
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  521 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-539-666-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 521;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    167 EAGGSS 172


RESULT  31
US-08-348-891A-2
; Sequence 2, Application US/08348891A
; Patent No. 5654136
;  GENERAL INFORMATION:
;     APPLICANT:  SASAKI, Keiko
;     APPLICANT:  MORI, Takayuki
;     APPLICANT:  MAKINO, Satoshi
;     TITLE OF INVENTION:  ATTENUATED MEASLES VIRUS VACCINE,
;     TITLE OF INVENTION:  CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
;     TITLE OF INVENTION:  ITS ABSOLUTE IDENTIFICATION
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  YOUNG & THOMPSON
;       STREET:  745 South 23rd Street
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  USA
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/348,891A
;       FILING DATE:  25-NOV-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/848,400
;       FILING DATE:  10-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 3-293625
;       FILING DATE:  14-OCT-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  KP-7501
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-521-2297
;       TELEFAX:  703-685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  2:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  525 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-348-891A-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 525;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     263 EAGLAS 268


RESULT  32
US-08-905-817-2
; Sequence 2, Application US/08905817
; Patent No. 5824777
;  GENERAL INFORMATION:
;    APPLICANT:  SASAKI, Keiko
;    APPLICANT:  MORI, Takayuki
;    APPLICANT:  MAKINO, Satoshi
;    TITLE OF INVENTION:  ATTENUATED MEASLES VIRUS VACCINE,
;    TITLE OF INVENTION:  CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
;    TITLE OF INVENTION:  ITS ABSOLUTE IDENTIFICATION
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,817
;      FILING DATE:  04-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/348,891
;      FILING DATE:  25-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/848,400
;      FILING DATE:  10-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 3-293625
;      FILING DATE:  14-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  KP-7501A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-521-2297
;      TELEFAX:  703-685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  525 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-905-817-2


  Query Match             85.7%;  Score 18;  DB 2;  Length 525;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     263 EAGLAS 268


RESULT  33
US-08-513-841-1
; Sequence 1, Application US/08513841
; Patent No. 5753481
;  GENERAL INFORMATION:
;    APPLICANT:  Niwa, Mineo
;    APPLICANT:  Saito, Yoshimasa
;    APPLICANT:  Ishii, Yoshinori
;    APPLICANT:  Yoshida, Masaru
;    APPLICANT:  Suzuki, Hiromi
;    TITLE OF INVENTION:  No. 5753481el L-sorbose Dehydrogenase and No. 5753481el L-s
;    TITLE OF INVENTION:  Dehydrogenase Obtained from Gluconobacter oxydans T-100
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS-DOS Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/513,841
;      FILING DATE:  01-NOV-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  UK 9304700.9
;      FILING DATE:  08-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 241851/1993
;      FILING DATE:  28-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORMAN F. OBLON
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-909-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Gluconobacter oxydans
;      STRAIN:  T-100
;    FEATURE:
;      NAME/KEY:  mat peptide
;      LOCATION:  1..530
;      IDENTIFICATION METHOD:  experimentally
US-08-513-841-1


  Query Match             85.7%;  Score 18;  DB 1;  Length 530;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     364 EAGVTS 369
```

```
RESULT  34
US-08-696-834-1
; Sequence 1, Application US/08696834
; Patent No. 5834263
;  GENERAL INFORMATION:
;    APPLICANT:  Niwa, Mineo
;    APPLICANT:  Saito, Yoshimasa
;    APPLICANT:  Ishii, Yoshinori
;    APPLICANT:  Yoshida, Masaru
;    APPLICANT:  Hayashi, Hiromi
;    TITLE OF INVENTION:  Method for Producing 2-Keto-L-Gulonic Acid
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/696,834
;      FILING DATE:  24-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 28612/1994
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Gluconobacter oxydans
;      STRAIN:  T-100
;    FEATURE:
;      NAME/KEY:  mat peptide
;      LOCATION:  1..530
;      IDENTIFICATION METHOD:  experimentally
US-08-696-834-1


  Query Match             85.7%;  Score 18;  DB 2;  Length 530;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     364 EAGVTS 369


RESULT  35
US-08-942-673-1
; Sequence 1, Application US/08942673
; Patent No. 5861292
;  GENERAL INFORMATION:
;    APPLICANT:  Niwa, Mineo
;    APPLICANT:  Saito, Yoshimasa
;    APPLICANT:  Ishii, Yoshinori
;    APPLICANT:  Yoshida, Masaru
;    APPLICANT:  Suzuki, Hiromi
;    TITLE OF INVENTION:  No. 5861292el L-sorbose Dehydrogenase and No. 5861292el
;    TITLE OF INVENTION:  L-sorbosone Dehydrogenase Obtained from Gluconobacter
;    TITLE OF INVENTION:  oxydans T-100
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS-DOS Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/942,673
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/513,841
;      FILING DATE:  01-NOV-1995
;      APPLICATION NUMBER:  UK 9304700.9
;      FILING DATE:  08-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 241851/1993
;      FILING DATE:  28-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORMAN F. OBLON
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-909-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Gluconobacter oxydans
;      STRAIN:  T-100
;    FEATURE:
;      NAME/KEY:  mat peptide
;      LOCATION:  1..530
;      IDENTIFICATION METHOD:  experimentally
US-08-942-673-1


  Query Match             85.7%;  Score 18;  DB 2;  Length 530;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     364 EAGVTS 369


RESULT  36
US-09-118-317-1
; Sequence 1, Application US/09118317
; Patent No. 6197562
;  GENERAL INFORMATION:
;    APPLICANT:  Niwa, Mineo
;    APPLICANT:  Saito, Yoshimasa
;    APPLICANT:  Ishii, Yoshinori
```

```
;    APPLICANT:  Yoshida, Masaru
;    APPLICANT:  Suzuki, Hiromi
;    TITLE OF INVENTION:  No. 6197562el L-sorbose Dehydrogenase and No. 6197562el
;    TITLE OF INVENTION:  L-sorbosone Dehydrogenase Obtained from Gluconobacter
;    TITLE OF INVENTION:  oxydans T-100
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS-DOS Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/118,317
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/513,841
;      FILING DATE:  01-NOV-1995
;      APPLICATION NUMBER:  UK 9304700.9
;      FILING DATE:  08-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 241851/1993
;      FILING DATE:  28-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORMAN F. OBLON
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  18-909-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Gluconobacter oxydans
;      STRAIN:  T-100
;    FEATURE:
;      NAME/KEY:  mat peptide
;      LOCATION:  1..530
;      IDENTIFICATION METHOD:  experimentally
US-09-118-317-1


  Query Match             85.7%;  Score 18;  DB 4;  Length 530;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    364 EAGVTS 369


RESULT  37
US-08-808-931-18
; Sequence 18, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  560 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-931-18


  Query Match             85.7%;  Score 18;  DB 2;  Length 560;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     67 EAGSGS 72


RESULT  38
US-08-808-323-18
; Sequence 18, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  560 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-323-18


  Query Match             85.7%;  Score 18;  DB 3;  Length 560;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
            |||  |
Db      67 EAGSGS 72


RESULT  39
US-09-050-603A-18
; Sequence 18, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  560 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-09-050-603A-18


  Query Match             85.7%;  Score 18;  DB 3;  Length 560;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
            |||  |
Db      67 EAGSGS 72


RESULT  40
US-09-102-420B-18
; Sequence 18, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  560 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-09-102-420B-18


  Query Match             85.7%;  Score 18;  DB 3;  Length 560;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      67 EAGSGS 72


RESULT  41
US-09-497-698-18
; Sequence 18, Application US/09497698
; Patent No. 6308458
;   GENERAL INFORMATION:
;        APPLICANT: Volrath, Sandra
;                   Johnson, Marie
;                   Ward, Eric
;                   Heifetz, Peter
;        TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;                            OXIDASE ("PROTOX")
;        NUMBER OF SEQUENCES: 43
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: No. 6308458artis Corporation
;             STREET: 3054 Cornwallis Road
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/497,698
;             FILING DATE: 03-Feb-2000
;             CLASSIFICATION: <Unknown>
;                             30-MAR-1998
;                             11-MAR-1998
;                             28-FEB-1997
;                             28-FEB-1996
;                             28-FEB-1996
;                             21-JUN-1996
;                             06-JUN-1995
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/102,420
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 09/050,603
;             FILING DATE: 30-MAR-1998
;             APPLICATION NUMBER: US 60/126,430
;             FILING DATE: 11-MAR-1998
;             APPLICATION NUMBER: US 08/808,931
;             FILING DATE: 28-FEB-1997
;             APPLICATION NUMBER: US 60/012,705
;             FILING DATE: 28-FEB-1996
;             APPLICATION NUMBER: US 60/013,612
;             FILING DATE: 28-FEB-1996
;             APPLICATION NUMBER: US 60/020,003
;             FILING DATE: 21-JUN-1996
;             APPLICATION NUMBER: US 08/472,028
;             FILING DATE: 06-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Meigs, J. Timothy
;             REGISTRATION NUMBER: 38,241
;             REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (919) 541-8587
;             TELEFAX: (919) 541-8689
;   INFORMATION FOR SEQ ID NO: 18:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 560 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: No. 6308458 Relevant
;             TOPOLOGY: No. 6308458 Relevant
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-497-698-18


  Query Match             85.7%;  Score 18;  DB 4;  Length 560;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      67 EAGSGS 72


RESULT  42
US-08-419-078-2
; Sequence 2, Application US/08419078
; Patent No. 5587306
;  GENERAL INFORMATION:
;    APPLICANT:  HAWKINS, PHILLIP R.
;    APPLICANT:  SEILHAMER, JEFFREY J.
;    TITLE OF INVENTION:  PHOSPHOLIPASE C HOMOLOG
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3330 HILLVIEW AVENUE
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/419,078
;      FILING DATE:
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LUTHER, BARBARA J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PF0030 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-855-0572
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  566 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      LIBRARY:  No. 5587306e
;      CLONE:  9118
US-08-419-078-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 566;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    338 EAGAQS 343


RESULT  43
US-08-726-883-2
; Sequence 2, Application US/08726883
; Patent No. 5676946
;  GENERAL INFORMATION:
;    APPLICANT:  HAWKINS, PHILLIP R.
;    APPLICANT:  SEILHAMER, JEFFREY J.
;    TITLE OF INVENTION:  PHOSPHOLIPASE C HOMOLOG
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3330 HILLVIEW AVENUE
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,883
;      FILING DATE:  04-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/419,078
;      FILING DATE:  10-APR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LUTHER, BARBARA J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PF0030 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-855-0572
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  566 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      LIBRARY:  No. 5676946e
;      CLONE:  9118
US-08-726-883-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 566;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    338 EAGAQS 343


RESULT  44
US-08-696-944-2
; Sequence 2, Application US/08696944
; Patent No. 5981831
;  GENERAL INFORMATION:
;    APPLICANT:  Sumant CHENGAPPA
;    APPLICANT:  Susan A. HELLYER
;    APPLICANT:  John S. REID
;    APPLICANT:  Jacqueline DE SILVA
;    TITLE OF INVENTION:  No. 5981831el Exo-(1-4)-Beta-D Galactanase
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/696,944
;      FILING DATE:  23-AUG-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/00372
;      FILING DATE:  23-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9403423.8
;      FILING DATE:  23-FEB-1994
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  730 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-696-944-2


  Query Match             85.7%;  Score 18;  DB 2;  Length 730;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    603 EAGSNS 608


RESULT  45
US-08-731-716-2
; Sequence 2, Application US/08731716
; Patent No. 5789202
;  GENERAL INFORMATION:
;    APPLICANT:  Hoskins, JoAnn
;    APPLICANT:  Jaskunas, S. Richard
;    APPLICANT:  Rockey, Pamela K.
;    APPLICANT:  Zhao, Genshi
```

```
;    APPLICANT:  Rosteck, Paul R. Jr.
;    APPLICANT:  No. 5789202ris, Franklin H.
;    TITLE OF INVENTION:  Penicillin Binding Protein From
;    TITLE OF INVENTION:  Streptococcus Pneumoniae
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/731,716
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-10,887
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  731 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-731-716-2


  Query Match             85.7%;  Score 18;  DB 1;  Length 731;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    102 EAGALS 107


RESULT  46
US-09-651-656-19
; Sequence 19, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
;  APPLICANT:  McCUTHEN-MALONEY, SANDRA
;  APPLICANT:  LAWRENCE LIVERMORE NATIONAL LABORATORY
;  TITLE OF INVENTION:  DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
;  TITLE OF INVENTION:  POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
;  TITLE OF INVENTION:  DNA DAMAGE AND DNA MISMATCHES
;  FILE REFERENCE:  IL-10689
;  CURRENT APPLICATION NUMBER: US/09/651,656
;  CURRENT FILING DATE:  2000-08-29
;  PRIOR APPLICATION NUMBER:  60/192,764
;  PRIOR FILING DATE:  2000-03-28
;  NUMBER OF SEQ ID NOS:  106
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 823
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-651-656-19


  Query Match             85.7%;  Score 18;  DB 4;  Length 823;
  Best Local Similarity   66.7%;  Pred. No. 3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    321 EAGSGS 326


RESULT  47
US-08-434-730-14
; Sequence 14, Application US/08434730
; Patent No. 5637463
;  GENERAL INFORMATION:
;    APPLICANT:  Dalton, Stephen
;    APPLICANT:  Kochan, Jarema P
;    APPLICANT:  Osborne, Mark A
;    TITLE OF INVENTION:  METHOD TO DETECT PROTEIN-PROTEIN
;    TITLE OF INVENTION:  INTERACTIONS
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/434,730
;      FILING DATE:  04-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Semionow, Raina
;      REGISTRATION NUMBER:  39022
;      REFERENCE/DOCKET NUMBER:  9069
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201)235-4391
;      TELEFAX:  (201)235-2363
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  968 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-434-730-14


  Query Match             85.7%;  Score 18;  DB 1;  Length 968;
  Best Local Similarity   66.7%;  Pred. No. 3.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    711 EAGVTS 716


RESULT  48
US-08-560-005-2
; Sequence 2, Application US/08560005
; Patent No. 6001354
;  GENERAL INFORMATION:
;    APPLICANT:  Pot, David A.
;    APPLICANT:  Williams, Lewis T.
;    APPLICANT:  Jefferson, Anne Bennett
;    APPLICANT:  Majerus, Philip W.
;    TITLE OF INVENTION:  No. 6001354el Grb2 Associating Protein and Nucleic
;    TITLE OF INVENTION:  Acids Encoding Therefor
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
```

```
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/560,005
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  2307K-0624000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  976 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-560-005-2


  Query Match             85.7%;  Score 18;  DB 3;  Length 976;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     496 EAGVTS 501


RESULT  49
US-09-195-868-14
; Sequence 14, Application US/09195868
; Patent No. 6090621
;  GENERAL INFORMATION:
;    APPLICANT:  KAVANAUGH MD, MICHAEL
;    APPLICANT:  POT PH.D., DAVID
;    APPLICANT:  WILLIAMS MDPHD, LEWIS T.
;    TITLE OF INVENTION:  SIGNALING INOSITOL POLYPHOSPHATE
;    TITLE OF INVENTION:  5-PHOSPHATASES (SIPs)
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHIRON CORPORATION
;      STREET:  4560 HORTON STREET
;      CITY:  EMERYVILLE
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/195,868
;      FILING DATE:
;      CLASSIFICATION:
;      APPLICATION NUMBER:  US/09/195,868
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FIRESTONE, LEIGH H.
;      REGISTRATION NUMBER:  36,831
;      REFERENCE/DOCKET NUMBER:  1182.004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-923-2707
;      TELEFAX:  510-655-3542
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  976 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-195-868-14


  Query Match             85.7%;  Score 18;  DB 3;  Length 976;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     496 EAGVTS 501


RESULT  50
US-09-418-540-2
; Sequence 2, Application US/09418540
; Patent No. 6296848
;  GENERAL INFORMATION:
;    APPLICANT:  Pot, David A.
;    APPLICANT:  Williams, Lewis T.
;    APPLICANT:  Jefferson, Anne Bennett
;    APPLICANT:  Majerus, Philip W.
;    TITLE OF INVENTION:  No. 6296848el Grb2 Associating Protein and Nucleic
;    TITLE OF INVENTION:  Acids Encoding Therefor
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/418,540
;      FILING DATE:  14-OCT-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/560,005
;      FILING DATE:  17-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  2307K-0624000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  976 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-418-540-2


  Query Match             85.7%;  Score 18;  DB 4;  Length 976;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
```

Db 496 EAGVTS 501

Search completed: August 30, 2002, 15:04:45
Job time: 8556 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 30, 2002, 15:00:04 ; Search time 26.93 Seconds
(without alignments)
21.409 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	90.5	37	2	A48620	adhesin - Staphylo
2	19	90.5	83	2	AF2563	hypothetical prote
3	19	90.5	122	2	C84320	hypothetical prote
4	19	90.5	143	2	B72627	hypothetical prote
5	19	90.5	145	2	F82189	hypothetical prote
6	19	90.5	164	2	T11215	hypothetical prote
7	19	90.5	200	2	T29807	hypothetical prote
8	19	90.5	201	2	T23855	hypothetical prote
9	19	90.5	204	2	A70844	probable moaE3 pro
10	19	90.5	212	2	AD1560	two-component resp
11	19	90.5	212	2	AF1202	two-component resp
12	19	90.5	220	2	AH0459	Sec-independent pr
13	19	90.5	245	2	E84169	hypothetical prote
14	19	90.5	246	2	AE1029	probable exported
15	19	90.5	257	2	T48058	RING-H2 zinc finge
16	19	90.5	260	2	B90026	hypothetical prote
17	19	90.5	263	2	G87721	protein ZC123.3 [i
18	19	90.5	266	2	JC1071	coat protein - soy
19	19	90.5	267	2	S18931	coat protein - soy
20	19	90.5	302	2	H82638	hypothetical prote
21	19	90.5	303	1	D64070	ATP phosphoribosyl
22	19	90.5	311	2	A56235	transcription acti
23	19	90.5	322	2	S38091	hypothetical prote
24	19	90.5	323	2	I49529	transcription fact
25	19	90.5	332	2	A38873	myristylated alani
26	19	90.5	352	2	S16547	neutral proteinase
27	19	90.5	352	2	G95872	probable adenylate
28	19	90.5	357	2	T01571	hypothetical prote
29	19	90.5	363	2	A83177	probable N-acetylg
30	19	90.5	363	2	T47240	amino acid transpo
31	19	90.5	365	2	E83800	hypothetical prote
32	19	90.5	367	1	I40226	3-isopropylmalate
33	19	90.5	387	2	T25452	hypothetical prote
34	19	90.5	389	1	A39429	cAMP response elem
35	19	90.5	390	2	JT0334	acid proteinase (E
36	19	90.5	393	2	F72068	dihydrolipoamide s
37	19	90.5	393	2	C86556	dihydrolipoamide s
38	19	90.5	397	2	AD1559	glycine betaine AB
39	19	90.5	397	2	AF1201	glycine betaine AB
40	19	90.5	403	2	B72778	probable processin
41	19	90.5	443	2	A54813	cAMP receptor CAR4
42	19	90.5	444	2	S35783	glycoprotein gX -
43	19	90.5	458	2	T02804	hypothetical prote
44	19	90.5	462	2	D85438	hypothetical prote
45	19	90.5	469	2	C87629	major facilitator
46	19	90.5	474	2	D70853	hypothetical prote
47	19	90.5	482	2	S08384	nodulation protein
48	19	90.5	482	2	S10133	nodulation protein
49	19	90.5	490	2	T41039	probable transcrip
50	19	90.5	492	2	S54536	probable membrane

ALIGNMENTS

RESULT 1
A48620
adhesin - Staphylococcus aureus (fragment)
C;Species: Staphylococcus aureus
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C;Accession: A48620
R;Patti, J.M.; Boles, J.O.; Hook, M.
Biochemistry 32, 11428-11435, 1993
A;Title: Identification and biochemical characterization of the ligand binding domain
A;Reference number: A48620; MUID:94032261
A;Contents: FDA 574
A;Accession: A48620
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-37 <PAT>
A;Note: sequence extracted from NCBI backbone (NCBIP:138726)

Query Match 90.5%; Score 19; DB 2; Length 37;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 eagxxs 6
          |||  |
Db     16 EAGTSS 21
```

RESULT 2
AF2563
hypothetical protein asl8505 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AF2563
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2563
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-83 <KUR>
A;Cross-references: GB:AP003604; PIDN:BAB77424.1; PID:g17134868; GSPDB:GN00183
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl8505

A;Genome: plasmid

Query Match 90.5%; Score 19; DB 2; Length 83;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 78 EAGASS 83

RESULT 3
C84320
hypothetical protein Vng1678h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84320
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483
A;Accession: C84320
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-122 <STO>
A;Cross-references: GB:AE004437; NID:g10581147; PIDN:AAG19927.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1678H

Query Match 90.5%; Score 19; DB 2; Length 122;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 57 EAGASS 62

RESULT 4
B72627
hypothetical protein APE1474 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: B72627
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339
A;Accession: B72627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-143 <KAW>
A;Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80472.1; PID:d1044258; PID:g510
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1474

Query Match 90.5%; Score 19; DB 2; Length 143;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 85 EAGAAS 90

RESULT 5
F82189
hypothetical protein VC1536 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: F82189
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: F82189
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-145 <HEI>
A;Cross-references: GB:AE004231; GB:AE003852; NID:g9656027; PIDN:AAF94690.1; GSPDB:GN
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1536
A;Map position: 1

Query Match 90.5%; Score 19; DB 2; Length 145;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 114 EAGSTS 119

RESULT 6
T11215
hypothetical protein 5 - Streptomyces glaucescens
C;Species: Streptomyces glaucescens
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: T11215
R;Summers, R.G.; Ali, A.; Shen, B.; Wessel, W.A.; Hutchinson, C.R.
Biochemistry 34, 9389-9402, 1995
A;Title: Malonyl-coenzyme A:acyl carrier protein acyltransferase of Streptomyces glau
A;Reference number: Z17254; MUID:95352622
A;Accession: T11215
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-164 <SUM>
A;Cross-references: EMBL:L43074; NID:g870805; PID:g870810

Query Match 90.5%; Score 19; DB 2; Length 164;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 55 EAGTAS 60

RESULT 7
T29807
hypothetical protein C25A8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29807
R;Latreille, P.; Stellyes, L.
submitted to the EMBL Data Library, June 1996
A;Description: The sequence of C. elegans cosmid C25A8.
A;Reference number: Z20689
A;Accession: T29807
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-200 <LAT>
A;Cross-references: EMBL:U61958; PIDN:AAB03180.1; GSPDB:GN00022; CESP:C25A8.2

A;Experimental source: strain Bristol N2; clone C25A8
C;Genetics:
A;Gene: CESP:C25A8.2
A;Map position: 4
A;Introns: 173/3

Query Match 90.5%; Score 19; DB 2; Length 200;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 109 EAGAAS 114

RESULT 8
T23855
hypothetical protein R02D5.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23855
R;Matthews, L.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19808
A;Accession: T23855
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-201 <WIL>
A;Cross-references: EMBL:Z78015; PIDN:CAB01436.1; GSPDB:GN00023; CESP:R02D5.7
A;Experimental source: clone R02D5
C;Genetics:
A;Gene: CESP:R02D5.7
A;Map position: 5
A;Introns: 174/3

Query Match 90.5%; Score 19; DB 2; Length 201;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 110 EAGAAS 115

RESULT 9
A70844
probable moaE3 protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: A70844
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: A70844
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-204 <COL>
A;Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAA17094.1; PID:e125115
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: moaE3

Query Match 90.5%; Score 19; DB 2; Length 204;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 153 EAGTAS 158

RESULT 10
AD1560
two-component response regulator, in particular B. subtilis YvqC protein homolog lin1
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AD1560
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1560
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96252.1; PID:g16413480; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1021
C;Superfamily: regulatory protein comA; response regulator homology

Query Match 90.5%; Score 19; DB 2; Length 212;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 95 EAGASS 100

RESULT 11
AF1202
two-component response regulator, in particular B. subtilis YvqC protein homolog lmo1
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AF1202
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99100.1; PID:g16410424; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1022
C;Superfamily: regulatory protein comA; response regulator homology

Query Match 90.5%; Score 19; DB 2; Length 212;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 95 EAGASS 100

RESULT 12
AH0459
Sec-independent protein translocase protein TatB [imported] - Yersinia pestis (strain CO
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0459
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0459
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC93244.1; PID:g15981690; GSPDB:GN00175
C;Genetics:
A;Gene: tatB

Query Match 90.5%; Score 19; DB 2; Length 220;
Best Local Similarity 66.7%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 143 EAGTAS 148

RESULT 13
E84169
hypothetical protein pimT1 [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84169
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483
A;Accession: E84169
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-245 <STO>
A;Cross-references: GB:AE004437; NID:g10579741; PIDN:AAG18721.1; GSPDB:GN00138
C;Genetics:
A;Gene: pimT1

Query Match 90.5%; Score 19; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 151 EAGAAS 156

RESULT 14
AE1029
probable exported protein STY4558 [imported] - Salmonella enterica subsp. enterica serov
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AE1029
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A;Reference number: AB0502; PMID:11677608
A;Accession: AE1029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD09334.1; PID:g16505334; GSPDB:GN00176
C;Genetics:
A;Gene: STY4558

Query Match 90.5%; Score 19; DB 2; Length 246;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 93 EAGSAS 98

RESULT 15
T48058
RING-H2 zinc finger protein ATL5 - Arabidopsis thaliana
N;Alternate names: protein F26K9.120
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Dec-2000
C;Accession: T48058
R;Bloecker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salano
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24465
A;Accession: T48058
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-257 <BLO>
A;Cross-references: EMBL:AL162651
A;Experimental source: cultivar Columbia; BAC clone F26K9
C;Genetics:
A;Map position: 3
A;Note: F26K9.120
C;Superfamily: Arabidopsis hypothetical protein F19I3.22; RING finger homology
F;109-160/Domain: RING finger homology <RRN>

Query Match 90.5%; Score 19; DB 2; Length 257;
Best Local Similarity 66.7%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 184 EAGSSS 189

RESULT 16
B90026
hypothetical protein modA [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: B90026
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: B90026
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <KUR>
A;Cross-references: GB:BA000018; PID:g13702079; PIDN:BAB43371.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: modA
C;Superfamily: molybdate-binding periplasmic protein

```
  Query Match             90.5%;  Score 19;  DB 2;  Length 260;
  Best Local Similarity   66.7%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    227 EAGATS 232


RESULT  17
G87721
protein ZC123.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: G87721
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G87721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <STO>
A;Cross-references: GB:chr_I; PIDN:AAB97603.1; PID:g2804499; GSPDB:GN00019; CESP:ZC123.3
A;Note: contains similarity to C2H2-type zinc fingers
C;Genetics:
A;Gene: ZC123.3
A;Map position: 1


  Query Match             90.5%;  Score 19;  DB 2;  Length 263;
  Best Local Similarity   66.7%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     31 EAGASS 36


RESULT  18
JC1071
coat protein - soybean mosaic virus
C;Species: soybean mosaic virus, SbMV
C;Date: 02-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 17-Nov-2000
C;Accession: JC1071
R;Chu, R.Y.; Leng, X.H.; Bao, Y.M.; Pan, N.S.; Pu, Z.Q.
Acta Bot. Sin. 34, 523-528, 1992
A;Title: Amplification of soybean mosaic virus coat protein gene by polymerase chain rea
A;Reference number: JC1071
A;Accession: JC1071
A;Molecule type: DNA
A;Residues: 1-266 <CHU>
C;Superfamily: tobacco etch virus genome polyprotein


  Query Match             90.5%;  Score 19;  DB 2;  Length 266;
  Best Local Similarity   66.7%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     25 EAGTSS 30


RESULT  19
S18931
coat protein - soybean mosaic virus (fragment)
C;Species: soybean mosaic virus, SbMV
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C;Accession: S18931
R;Chu, R.
submitted to the EMBL Data Library, January 1992
A;Description: cDNA sequence of the gene encoding coat protein of SMV.
A;Reference number: S18931
A;Accession: S18931
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <CHU>
A;Cross-references: EMBL:X63771; NID:g61983; PIDN:CAA45307.1; PID:g61984
C;Superfamily: tobacco etch virus genome polyprotein


  Query Match             90.5%;  Score 19;  DB 2;  Length 267;
  Best Local Similarity   66.7%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     26 EAGTSS 31


RESULT  20
H82638
hypothetical protein XF1783 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: H82638
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <SIM>
A;Cross-references: GB:AE004000; GB:AE003849; NID:g9106850; PIDN:AAF84591.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1783


  Query Match             90.5%;  Score 19;  DB 2;  Length 302;
  Best Local Similarity   66.7%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    106 EAGTAS 111


RESULT  21
D64070
ATP phosphoribosyltransferase (EC 2.4.2.17) - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000
C;Accession: D64070
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
```

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630
A;Accession: D64070
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-303 <TIGR>
A;Cross-references: GB:U32729; GB:L42023; NID:g1573439; PIDN:AAC22127.1; PID:g1573446; T
A;Note: named as homolog to a protein from Escherichia coli
C;Superfamily: ATP phosphoribosyltransferase; ATP phosphoribosyltransferase homology
C;Keywords: glycosyltransferase; histidine biosynthesis; pentosyltransferase

  Query Match 90.5%; Score 19; DB 1; Length 303;
  Best Local Similarity 66.7%; Pred. No. 6.6e+02;
  Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
Db 287 EAGASS 292

RESULT 22
A56235
transcription activator MafB - chicken
C;Species: Gallus gallus (chicken)
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
C;Accession: A56235
R;Kataoka, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M.
Mol. Cell. Biol. 14, 7581-7591, 1994
A;Title: MafB, a new Maf family transcription activator that can associate with Maf and
A;Reference number: A56235; MUID:95021288
A;Accession: A56235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <KAT>
A;Cross-references: GB:D28600; NID:g516723; PIDN:BAA05938.1; PID:g516724
C;Genetics:
A;Introns: #status absent
C;Superfamily: maf transforming protein; maf homology
C;Keywords: DNA binding; homodimer; leucine zipper
F;200-289/Domain: maf homology <MAF>

  Query Match 90.5%; Score 19; DB 2; Length 311;
  Best Local Similarity 66.7%; Pred. No. 6.7e+02;
  Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
Db 296 EAGSTS 301

RESULT 23
S38091
hypothetical protein YKR022c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999
C;Accession: S38091
R;Duesterhoeft, A.; Moestl, D.; Poehlmann, R.; Philippsen, P.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37811
A;Accession: S38091
A;Molecule type: DNA
A;Residues: 1-322 <DUE>
A;Cross-references: EMBL:Z28247; NID:g486444; PID:g486445; GSPDB:GN00011; MIPS:YKR022c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YKR022c
A;Map position: 11R

  Query Match 90.5%; Score 19; DB 2; Length 322;
  Best Local Similarity 66.7%; Pred. No. 7e+02;
  Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
Db 119 EAGSSS 124

RESULT 24
I49529
transcription factor-kr - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 16-Jul-1999
C;Accession: I49529
R;Cordes, S.P.; Barsh, G.S.
Cell 79, 1025-1034, 1994
A;Title: The mouse segmentation gene kr encodes a novel basic domain-leucine zipper t
A;Reference number: A55200; MUID:95094266
A;Accession: I49529
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-323 <RES>
A;Cross-references: GB:L36435; NID:g625043; PIDN:AAA65689.1; PID:g625044
C;Superfamily: maf transforming protein; maf homology
C;Keywords: leucine zipper; transcription factor
F;212-301/Domain: maf homology <MAF>

  Query Match 90.5%; Score 19; DB 2; Length 323;
  Best Local Similarity 66.7%; Pred. No. 7e+02;
  Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
 ||| |
Db 308 EAGSTS 313

RESULT 25
A38873
myristylated alanine-rich protein kinase C substrate - human
N;Alternate names: acidic calmodulin-binding 80K protein; MARCKS
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C;Accession: A38873; A42977; A40758; S29269
R;Shimizu, N.
submitted to DDBJ, September 1991
A;Reference number: A38873
A;Accession: A38873
A;Molecule type: mRNA
A;Residues: 1-332 <SHI>
A;Cross-references: GB:D10522; GB:D90498; NID:g219893; PIDN:BAA01392.1; PID:g219894
R;Sakai, K.; Hirai, M.; Kudoh, J.; Minoshima, S.; Shimizu, N.
Genomics 14, 175-178, 1992
A;Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L pro
A;Reference number: A42977; MUID:93052291
A;Accession: A42977
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-13,'G',15-332 <SAK>
A;Cross-references: GB:D90498
A;Experimental source: squamous carcinoma cells A431
A;Note: sequence extracted from NCBI backbone (NCBIP:118653)
R;Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.;
J. Biol. Chem. 266, 14399-14405, 1991
A;Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACS)
A;Reference number: A40758; MUID:91317795
A;Accession: A40758
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-118,'P',120-233,'W',235-286,'LVC',290,'RRGGSPRGGARGRRSLNQ',30
A;Cross-references: GB:M68956

A;Note: the authors translated the codon GGC for residue 53 as Arg
R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.
Eur. J. Biochem. 209, 7-14, 1992
A;Title: Relationship between the major protein kinase C substrates acidic 80-kDa protei
or equivalent genes in different species.
A;Reference number: S29267; MUID:93011168
A;Accession: S29269
A;Molecule type: mRNA
A;Residues: 189-223,'R',225-234,'E',236-322 <HER>
C;Comment: This protein is a major cellular substrate for protein kinase C and plays a r
C;Comment: It binds to calmodulin in one to one molar ratio in the presence of calcium a
C;Genetics:
A;Gene: GDB:MACS
A;Cross-references: GDB:118835; OMIM:177061
A;Map position: 6q22.2-6q22.2
C;Superfamily: neurofilament triplet H protein
C;Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyla
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;159,163,167,170/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

Query Match 90.5%; Score 19; DB 2; Length 332;
Best Local Similarity 66.7%; Pred. No. 7.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 206 EAGAAS 211

RESULT 26
S16547
neutral proteinase II - Aspergillus oryzae
C;Species: Aspergillus oryzae
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C;Accession: S16547; S47562
R;Tatsumi, H.; Murakami, S.; Tsuji, R.F.; Ishida, Y.; Murakami, K.; Masaki, A.; Kawabe,
Mol. Gen. Genet. 228, 97-103, 1991
A;Title: Cloning and expression in yeast of a cDNA clone encoding Aspergillus oryzae neu
A;Reference number: S16547; MUID:91360097
A;Accession: S16547
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-352 <TAT>
A;Cross-references: GB:S53810; NID:g234832; PIDN:AAB19701.1; PID:g234833
R;Tatsumi, H.; Ikegaya, K.; Murakami, S.; Kawabe, H.; Nakano, E.; Motai, H.
Biochim. Biophys. Acta 1208, 179-185, 1994
A;Title: Elucidation of the thermal stability of the neutral proteinase II from Aspergil
A;Reference number: S47562; MUID:94368822
A;Accession: S47562
A;Status: preliminary
A;Molecule type: protein
A;Residues: 176-352 <TA2>
C;Superfamily: Penicillium citrinum penicillolysin

Query Match 90.5%; Score 19; DB 2; Length 352;
Best Local Similarity 66.7%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 240 EAGSTS 245

RESULT 27
G95872
probable adenylate cyclase protein [imported] - Sinorhizobium meliloti (strain 1021) mag
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: G95872
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: G95872
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC48647.1; PID:g15140119; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20257
A;Genome: plasmid

Query Match 90.5%; Score 19; DB 2; Length 352;
Best Local Similarity 66.7%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 13 EAGTSS 18

RESULT 28
T01571
hypothetical protein A_TM018A10.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Nov-1999
C;Accession: T01571
R;Dempsey, S.; Harper, M.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of A. thaliana TM018A10.
A;Reference number: Z14348
A;Accession: T01571
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <DEM>
A;Cross-references: EMBL:AF013294; NID:g2252848; PID:g2252871
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 47/3; 201/2; 243/1; 259/2
A;Note: A_TM018A10.10
C;Superfamily: Arabidopsis thaliana hypothetical protein A_TM018A10.10

Query Match 90.5%; Score 19; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 7.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 91 EAGSSS 96

RESULT 29
A83177
probable N-acetylglucosamine-6-phosphate deacetylase PA3758 [imported] - Pseudomonas
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: A83177
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337
A;Accession: A83177
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-363 <STO>
A;Cross-references: GB:AE004794; GB:AE004091; NID:g9949917; PIDN:AAG07145.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3758

Query Match 90.5%; Score 19; DB 2; Length 363;
Best Local Similarity 66.7%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 196 EAGASS 201

RESULT 30
T47240
amino acid transport protein arg-1, mitochondrial [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C;Accession: T47240
R;Liu, Q.; Dunlap, J.C.
Genetics 143, 1163-1174, 1996
A;Title: Isolation and analysis of the arg-13 gene of Neurospora crassa.
A;Reference number: Z24416; MUID:96400914
A;Accession: T47240
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-363 <LIU>
A;Cross-references: EMBL:L36378; NID:g773383; PIDN:AAC37500.1; PID:g773384
A;Experimental source: strain bdA; isolate 30-1
C;Genetics:
A;Gene: arg-13
A;Map position: V
A;Introns: 50/3
C;Keywords: amino acid transport; mitochondrion

Query Match 90.5%; Score 19; DB 2; Length 363;
Best Local Similarity 66.7%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 14 EAGAAS 19

RESULT 31
E83800
hypothetical protein BH1205 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: E83800
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: E83800
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <STO>
A;Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04924.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1205

Query Match 90.5%; Score 19; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 247 EAGTAS 252

RESULT 32
I40226
3-isopropylmalate dehydrogenase (EC 1.1.1.85) - Bacillus megaterium
C;Species: Bacillus megaterium
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: I40226; S38506
R;Meinhardt, F.; Wittchen, K.D.; Buakamp, M.
Appl. Microbiol. Biotechnol. 41, 344-351, 1994
A;Title: Cloning and sequencing of the lenC and npM genes and a putative spoIV gene f
A;Reference number: I40226; MUID:94288995
A;Accession: I40226
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-367 <RES>
A;Cross-references: EMBL:X65184; NID:g414096; PIDN:CAA46295.1; PID:g414097
A;Experimental source: DSM 319
C;Genetics:
A;Gene: leuC
C;Function:
A;Pathway: leucine biosynthesis
C;Superfamily: 3-isopropylmalate dehydrogenase
C;Keywords: leucine biosynthesis; NAD; oxidoreductase

Query Match 90.5%; Score 19; DB 1; Length 367;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 48 EAGSSS 53

RESULT 33
T25452
hypothetical protein B0412.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25452
R;Bentley, D.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid B0412.
A;Reference number: Z20037
A;Accession: T25452
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-387 <BEN>
A;Cross-references: EMBL:U80953; PIDN:AAB52555.1; GSPDB:GN00021; CESP:B0412.1
A;Experimental source: strain Bristol N2; clone B0412
C;Genetics:
A;Gene: CESP:B0412.1
A;Map position: 3
A;Introns: 110/3; 146/2; 175/1; 213/2; 253/3; 318/1

Query Match 90.5%; Score 19; DB 2; Length 387;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 276 EAGSSS 281

RESULT 34
A39429
cAMP response element-binding protein ATF2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A39429
R;Kageyama, R.; Sasai, Y.; Nakanishi, S.
J. Biol. Chem. 266, 15525-15531, 1991
A;Title: Molecular characterization of transcription factors that bind to the cAMP respo
A;Reference number: A39429; MUID:91332085
A;Accession: A39429
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <KAG>
A;Cross-references: GB:M65148; NID:g206569; PIDN:AAA42013.1; PID:g206570
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;231-271/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 90.5%; Score 19; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 354 EAGATS 359

RESULT 35
JT0334
acid proteinase (EC 3.4.23.-) PEPI precursor - yeast (Saccharomycopsis fibuligera)
C;Species: Saccharomycopsis fibuligera
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Sep-1999
C;Accession: JT0334
R;Hirata, D.; Fukui, S.; Yamashita, I.
Agric. Biol. Chem. 52, 2647-2649, 1988
A;Title: Nucleotide sequence of the secretable acid protease gene PEPI in the yeast sacc
A;Reference number: JT0334
A;Accession: JT0334
A;Molecule type: DNA
A;Residues: 1-390 <HIR>
C;Genetics:
A;Gene: PEP1
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; hydrolase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-390/Product: acid proteinase #status predicted <MAT>
F;93,282/Active site: Asp #status predicted

Query Match 90.5%; Score 19; DB 2; Length 390;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 266 EAGSSS 271

RESULT 36
F72068
dihydrolipoamide succinyltransferase - Chlamydophila pneumoniae (strain CWL029)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: F72068
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: F72068
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-393 <ARN>
A;Cross-references: GB:AE001637; GB:AE001363; NID:g4376807; PIDN:AAD18667.1; PID:g437
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: sucB_2
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 90.5%; Score 19; DB 2; Length 393;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 101 EAGSSS 106

RESULT 37
C86556
dihydrolipoamide succinyltransferase [imported] - Chlamydophila pneumoniae (strain J1
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: C86556
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: C86556
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-393 <STO>
A;Cross-references: GB:BA000008; NID:g8978898; PIDN:BAA98733.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: sucB_2
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 90.5%; Score 19; DB 2; Length 393;
Best Local Similarity 66.7%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 101 EAGSSS 106

RESULT 38
AD1559
glycine betaine ABC transporter (ATP-binding protein) homolog gbuA [imported] - Liste
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AD1559
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96244.1; PID:g16413472; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: gbuA
C;Superfamily: glycine betaine/proline transport protein proV; ATP-binding cassette h

Query Match 90.5%; Score 19; DB 2; Length 397;

```
  Best Local Similarity   66.7%;  Pred. No. 8.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    301 EAGTSS 306


RESULT  39
AF1201
glycine betaine ABC transporter (ATP-binding protein) homolog gbuA [imported] - Listeria
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AF1201
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99092.1; PID:g16410416; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: gbuA
C;Superfamily: glycine betaine/proline transport protein proV; ATP-binding cassette homc


  Query Match            90.5%;  Score 19;  DB 2;  Length 397;
  Best Local Similarity   66.7%;  Pred. No. 8.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    301 EAGTSS 306


RESULT  40
B72778
probable processing proteinase APE0212 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: B72778
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339
A;Accession: B72778
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <KAW>
A;Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79124.1; PID:g5103603
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0212
C;Superfamily: mitochondrial processing peptidase alpha chain


  Query Match            90.5%;  Score 19;  DB 2;  Length 403;
  Best Local Similarity   66.7%;  Pred. No. 8.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    266 EAGATS 271


RESULT  41
A54813
cAMP receptor CAR4 - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 07-May-1999
C;Accession: A54813
R;Louis, J.M.; Ginsburg, G.T.; Kimmel, A.R.
Genes Dev. 8, 2086-2096, 1994
A;Title: The cAMP receptor CAR4 regulates axial patterning and cellular differentiati
A;Reference number: A54813; MUID:95047357
A;Accession: A54813
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A;Molecule type: DNA
A;Residues: 1-443 <LOU>
C;Genetics:
A;Gene: CAR4
C;Keywords: cAMP binding


  Query Match            90.5%;  Score 19;  DB 2;  Length 443;
  Best Local Similarity   66.7%;  Pred. No. 9.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     72 EAGSTS 77


RESULT  42
S35783
glycoprotein gX - bovine herpesvirus 1
C;Species: bovine herpesvirus 1
C;Date: 09-Jun-1994 #sequence_revision 12-May-1995 #text_change 08-Oct-1999
C;Accession: S35783
R;Audonnet, J.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35782
A;Accession: S35783
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <AUD>
A;Cross-references: EMBL:Z23068; NID:g312185; PIDN:CAA80603.1; PID:g312187
C;Superfamily: pseudorabies virus glycoprotein gX
C;Keywords: glycoprotein


  Query Match            90.5%;  Score 19;  DB 2;  Length 444;
  Best Local Similarity   66.7%;  Pred. No. 9.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    267 EAGSAS 272


RESULT  43
T02804
hypothetical protein L2602.2 [imported] - Leishmania major (strain Friedlin)
C;Species: Leishmania major
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C;Accession: H81456; T02804
R;Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protei
A;Reference number: A81455; MUID:99178987
A;Accession: H81456
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-458 <PYL>
A;Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24627.1; PID:g2995580; GSPDB:G
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L2602.2
```

```
A;Map position: 1
C;Superfamily: Leishmania major hypothetical protein L2602.2


  Query Match             90.5%;  Score 19;  DB 2;  Length 458;
  Best Local Similarity   66.7%;  Pred. No. 9.7e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     10 EAGTAS 15


RESULT  44
D85438
hypothetical protein AT4g37110 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: D85438
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488
A;Accession: D85438
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <STO>
A;Cross-references: GB:NC_001268; NID:g7270660; PIDN:CAB80377.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g37110
A;Map position: 4


  Query Match             90.5%;  Score 19;  DB 2;  Length 462;
  Best Local Similarity   66.7%;  Pred. No. 9.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    108 EAGAAS 113


RESULT  45
C87629
major facilitator family transporter CC3069 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: C87629
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-469 <STO>
A;Cross-references: GB:AE005673; NID:g13424719; PIDN:AAK25031.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3069
C;Superfamily: lincomycin-resistance protein lmrB


  Query Match             90.5%;  Score 19;  DB 2;  Length 469;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    394 EAGAAS 399


RESULT  46
D70853
hypothetical protein Rv3088 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70853
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987
A;Accession: D70853
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-474 <COL>
A;Cross-references: GB:AL021309; GB:AL123456; NID:g3261510; PIDN:CAA16146.1; PID:e124
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv3088


  Query Match             90.5%;  Score 19;  DB 2;  Length 474;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    460 EAGTTS 465


RESULT  47
S08384
nodulation protein nodT - Rhizobium leguminosarum plasmid pIJ1089
C;Species: Rhizobium leguminosarum
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 18-Sep-1998
C;Accession: S08384
R;Economou, A.; Hamilton, W.D.O.; Johnston, A.W.B.; Downie, J.A.
EMBO J. 9, 349-354, 1990
A;Title: The Rhizobium nodulation gene nodO encodes a Ca2+-binding protein that is ex
A;Reference number: S08384; MUID:90151607
A;Accession: S08384
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-482 <ECO>
A;Cross-references: EMBL:X17285
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1989
C;Genetics:
A;Genome: plasmid pIJ1089
C;Superfamily: nodulation protein nodT


  Query Match             90.5%;  Score 19;  DB 2;  Length 482;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    201 EAGAAS 206


RESULT  48
S10133
nodulation protein nodT - Rhizobium leguminosarum bv. viciae
C;Species: Rhizobium leguminosarum bv. viciae
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C;Accession: S10133
R;Surin, B.P.; Watson, J.M.; Hamilton, W.D.O.; Economou, A.; Downie, J.A.
Mol. Microbiol. 4, 245-252, 1990
A;Title: Molecular characterization of the nodulation gene, nodT, from two biovars of
A;Reference number: S08616; MUID:90251164
A;Accession: S10133
```

A;Molecule type: DNA
A;Residues: 1-482 <SUR>
A;Cross-references: EMBL:X17285; NID:g46251; PIDN:CAA35177.1; PID:g581512
C;Genetics:
A;Gene: nodT
A;Start codon: TTG
C;Superfamily: nodulation protein nodT

Query Match 90.5%; Score 19; DB 2; Length 482;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 201 EAGAAS 206

RESULT 49
T41039
probable transcription initiation factor Iif, alpha subunit - fission yeast (Schizosacch
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
C;Accession: T41039
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21966
A;Accession: T41039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-490 <LYN>
A;Cross-references: EMBL:AL034491; PIDN:CAA22493.1; GSPDB:GN00068; SPDB:SPCC1620.09c
A;Experimental source: strain 972h-; cosmid c1620
C;Genetics:
A;Gene: SPDB:SPCC1620.09c
A;Map position: 3
A;Introns: 180/3; 212/3
C;Keywords: transcription initiation

Query Match 90.5%; Score 19; DB 2; Length 490;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 234 EAGSAS 239

RESULT 50
S54536
probable membrane protein YDR240c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YD8419.07c
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 04-Mar-2000
C;Accession: S54536
R;Oliver, K.; Harris, D.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54530
A;Accession: S54536
A;Molecule type: DNA
A;Residues: 1-492 <OLI>
A;Cross-references: EMBL:Z49701; NID:g817819; PID:g817826; GSPDB:GN00004; MIPS:YDR240c
A;Experimental source: strain AB972
C;Genetics:
A;Gene: MIPS:YDR240c
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YDR240c
C;Keywords: transmembrane protein
F;119-135/Domain: transmembrane #status predicted <TMM>

Query Match 90.5%; Score 19; DB 2; Length 492;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6
||| |
Db 325 EAGATS 330

Search completed: August 30, 2002, 15:09:49
Job time: 585 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 30, 2002, 15:06:44 ; Search time 13.51 Seconds
(without alignments)
17.196 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	19	90.5	303	1	HIS1_HAEIN	P43853	haemophilus
2	19	90.5	322	1	YK02_YEAST	P36118	saccharomyc
3	19	90.5	323	1	MAF1_MOUSE	P54841	mus musculu
4	19	90.5	323	1	MAF1_RAT	P54842	rattus norv
5	19	90.5	331	1	MACS_HUMAN	P29966	homo sapien
6	19	90.5	352	1	NPII_ASPOR	P46076	aspergillus
7	19	90.5	363	1	AR13_NEUCR	Q01356	neurospora
8	19	90.5	367	1	LEU3_BACME	P41019	bacillus me
9	19	90.5	390	1	CARP_SACFI	P22929	saccharomyc
10	19	90.5	443	1	CAR4_DICDI	Q9tx43	dictyosteli
11	19	90.5	444	1	VGLX_HSVBS	Q08103	bovine herp
12	19	90.5	474	1	YU88_MYCTU	O53305	mycobacteri
13	19	90.5	482	1	NODT_RHILV	P15727	rhizobium l
14	19	90.5	487	1	ATF2_RAT	Q00969	rattus norv
15	19	90.5	561	1	MERA_ACICA	Q52109	acinetobact
16	19	90.5	561	1	MERA_ENTAG	P94702	enterobacte
17	19	90.5	578	1	AC22_STRCO	P46105	streptomyce
18	19	90.5	636	1	DNK2_SYNY3	P22358	synechocyst
19	19	90.5	638	1	TOXA_PSEAE	P11439	pseudomonas
20	19	90.5	643	1	SGT1_ARATH	Q9lsm5	arabidopsis
21	19	90.5	692	1	EOMD_XENLA	P79944	xenopus lae
22	19	90.5	883	1	PGCB_MOUSE	Q61361	mus musculu
23	19	90.5	883	1	PGCB_RAT	P55068	rattus norv
24	19	90.5	908	1	SRCA_RABIT	P13666	oryctolagus
25	19	90.5	1012	1	POLS_IBDVA	P08364	avian infec
26	19	90.5	1021	1	TSCC_HUMAN	P55017	homo sapien
27	19	90.5	1027	1	ISWI_DROME	Q24368	drosophila
28	19	90.5	1058	1	PMA1_DICDI	P54679	dictyosteli
29	19	90.5	1069	1	ACAA_ARATH	Q9szr1	arabidopsis
30	19	90.5	1183	1	CNA_STAAU	Q53654	staphylococ
31	19	90.5	1237	1	B3A2_MOUSE	P13808	mus musculu
32	19	90.5	1262	1	MYO6_HUMAN	Q9um54	homo sapien
33	19	90.5	1265	1	DYNA_DROME	P13496	drosophila
34	19	90.5	1265	1	MYO6_MOUSE	Q64331	mus musculu
35	19	90.5	1278	1	Y232_HUMAN	Q92628	homo sapien
36	19	90.5	1395	1	CUT1_MOUSE	P53564	mus musculu
37	19	90.5	1486	1	MUKB_ECOLI	P22523	escherichia
38	19	90.5	1505	1	CUT1_HUMAN	P39880	homo sapien
39	19	90.5	1548	1	UGGG_DROME	Q09332	drosophila
40	19	90.5	2035	1	HMP2_YEREN	P48633	yersinia en
41	18	85.7	30	1	CLPA_PINPS	P81671	pinus pinas
42	18	85.7	99	1	PLAS_CAPBU	P00294	capsella bu
43	18	85.7	123	1	WN14_HUMAN	O14904	homo sapien
44	18	85.7	143	1	PPIB_BACSU	P35137	bacillus su
45	18	85.7	147	1	R157_BOVIN	Q28183	bos taurus
46	18	85.7	154	1	DUT_MYCTU	O07199	mycobacteri
47	18	85.7	155	1	MOAE_RHIME	Q92qx5	rhizobium m
48	18	85.7	156	1	CEST_ECO57	P58233	escherichia
49	18	85.7	157	1	FLIN_PSEAE	Q51466	pseudomonas
50	18	85.7	159	1	YMH2_CAEEL	P34469	caenorhabdi

ALIGNMENTS

RESULT 1
HIS1_HAEIN
ID HIS1_HAEIN STANDARD; PRT; 303 AA.
AC P43853;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP phosphoribosyltransferase (EC 2.4.2.17).
GN HISG OR HI0468.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribosyl)-ATP + diphosphate =
CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: FIRST STEP IN HISTIDINE BIOSYNTHETIC PATHWAY. IS VERY
CC IMPORTANT IN THE REGULATION OF HISTIDINE METABOLISM.
CC -!- SUBUNIT: HOMOHEXAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ATP PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC --

CC --
DR EMBL; U32729; AAC22127.1; -.
DR TIGR; HI0468; -.
DR InterPro; IPR001348; HisG.
DR Pfam; PF01634; HisG; 1.
DR ProDom; PD003516; HisG; 1.
DR PROSITE; PS01316; ATP_P_PHORIBOSYLTR; 1.

```
KW   Histidine biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
SQ   SEQUENCE   303 AA;  33821 MW;  08C14D1F6E98A31D CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 303;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    287 EAGASS 292


RESULT   2
YK02_YEAST
ID   YK02_YEAST     STANDARD;      PRT;   322 AA.
AC   P36118;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hypothetical 36.6 kDa protein in YPT52-DBP7 intergenic region.
GN   YKR022C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   Duesterhoeft A., Moestl D., Poehlmann R., Philipssen P.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z28247; CAA82094.1; -.
DR   PIR; S38091; S38091.
DR   SGD; S0001730; YKR022C.
KW   Hypothetical protein.
SQ   SEQUENCE   322 AA;  36647 MW;  D7A601A46839244C CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 322;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    119 EAGSSS 124


RESULT   3
MAF1_MOUSE
ID   MAF1_MOUSE     STANDARD;      PRT;   323 AA.
AC   P54841;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transcription factor MAF1 (Segmentation protein KR) (Kreisler).
GN   MAFB OR MAF1 OR KRML.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95094266; PubMed=8001130;
RA   Cordes S.P., Barsh G.S.;
RT   "The mouse segmentation gene kr encodes a novel basic domain-leucine
RT   zipper transcription factor.";
RL   Cell 79:1025-1034(1994).
CC   -!- FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN
CC       SEGMENTATION).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN.
CC   -!- DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND
CC       IN THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES),
CC       THE HIGH LEVEL DOMAIN EXHIBITS A SHARP ROSTRAL EDGE COINCIDENT
CC       WITH THE R4/R5 BOUNDARY AND A DIFFUSE CAUDAL EDGE LOCATED MIDWAY
CC       THROUGH R6.
CC   -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L36435; AAA65689.1; -.
DR   HSSP; P05412; 1JUN.
DR   TRANSFAC; T01439; -.
DR   MGD; MGI:104555; Mafb.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN      131    143       POLY-HIS.
FT   DOMAIN      158    167       POLY-HIS.
FT   DNA_BIND    238    264       BASIC MOTIF.
FT   DOMAIN      266    287       LEUCINE-ZIPPER.
FT   MUTAGEN     248    248       N->S: LOSS OF TRANSCRIPTIONAL ACTIVITY.
SQ   SEQUENCE   323 AA;  35809 MW;  D77AE07ABD9C2AD2 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 323;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    308 EAGSTS 313


RESULT   4
MAF1_RAT
ID   MAF1_RAT       STANDARD;      PRT;   323 AA.
AC   P54842;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transcription factor MAF1.
GN   MAFB OR MAF1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=97190228; PubMed=9038383;
RA   Sakai M., Imaki J., Yoshida K., Ogata A., Matsushima-Hibaya Y.,
RA   Kuboki Y., Nishizawa M., Nishi S.;
RT   "Rat maf related genes: specific expression in chondrocytes, lens and
RT   spinal cord.";
RL   Oncogene 14:745-750(1997).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U56241; AAB50062.1; -.
DR   HSSP; P05412; 1JUN.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN      131    143       POLY-HIS.
FT   DOMAIN      158    167       POLY-HIS.
FT   DOMAIN      194    201       POLY-ALA.
FT   DNA_BIND    238    264       BASIC MOTIF.
FT   DOMAIN      266    287       LEUCINE-ZIPPER.
SQ   SEQUENCE   323 AA;  35792 MW;  6E386340D1F840A5 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 323;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    308 EAGSTS 313


RESULT   5
MACS_HUMAN
ID   MACS_HUMAN     STANDARD;      PRT;   331 AA.
AC   P29966;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myristoylated alanine-rich C-kinase substrate (MARCKS) (Protein kinase
DE   C substrate, 80 kDa protein, light chain) (PKCSL) (80K-L protein).
GN   MACS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91317795; PubMed=1860846;
RA   Harlan D.M., Graff J.M., Stumpo D.J., Eddy R.L. Jr., Shows T.B.,
RA   Boyle J.M., Blackshear P.J.;
RT   "The human myristoylated alanine-rich C kinase substrate (MARCKS)
RT   gene (MACS). Analysis of its gene product, promoter, and chromosomal
RT   localization.";
RL   J. Biol. Chem. 266:14399-14405(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93052291; PubMed=1427823;
RA   Sakai K., Hirai M., Kudoh J., Minoshima S., Shimizu N.;
RT   "Molecular cloning and chromosomal mapping of a cDNA encoding human
RT   80K-L protein: major substrate for protein kinase C.";
RL   Genomics 14:175-178(1992).
CC   -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC       PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC       SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC   -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC       ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M68956; AAA59555.1; -.
DR   EMBL; M68955; AAA59554.1; -.
DR   EMBL; D10522; BAA01392.1; -.
DR   PIR; A38873; A38873.
DR   MIM; 177061; -.
DR   InterPro; IPR002101; MARCKS.
DR   Pfam; PF02063; MARCKS; 1.
DR   PRINTS; PR00963; MARCKS.
DR   PROSITE; PS00826; MARCKS_1; 1.
DR   PROSITE; PS00827; MARCKS_2; 1.
KW   Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
KW   Membrane.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   LIPID         1      1       MYRISTATE (BY SIMILARITY).
FT   DOMAIN      151    175       CALMODULIN-BINDING (PSD).
FT   MOD_RES     158    158       PHOSPHORYLATION (BY PKC).
FT   MOD_RES     162    162       PHOSPHORYLATION (BY PKC).
FT   MOD_RES     166    166       PHOSPHORYLATION (BY PKC).
FT   MOD_RES     169    169       PHOSPHORYLATION (BY PKC).
FT   CONFLICT     83     83       S -> A (IN REF. 1).
FT   CONFLICT    118    118       A -> P (IN REF. 1).
FT   CONFLICT    233    233       P -> S (IN REF. 1).
FT   CONFLICT    286    307       PGAPPEQEAAPAEEPAAAAASS ->
FT                                LVCPRRGGSPRGGARGRRSLNQ (IN REF. 1).
SQ   SEQUENCE   331 AA;  31413 MW;  BCC837D586581774 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 331;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    205 EAGAAS 210


RESULT   6
NPII_ASPOR
ID   NPII_ASPOR     STANDARD;      PRT;   352 AA.
AC   P46076;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Neutral protease II precursor (EC 3.4.24.39) (Deuterolysin) (NPII).
OS   Aspergillus oryzae.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5062;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 176-210; 279-281 AND 304-341.
RC   STRAIN=ATCC 20386;
RX   MEDLINE=91360097; PubMed=1886621;
RA   Tatsumi H., Murakami S., Tsuji R.F., Ishida Y., Murakami K.,
RA   Masaki A., Kawabe H., Arimura H., Nakano E., Motai H.;
RT   "Cloning and expression in yeast of a cDNA clone encoding Aspergillus
RT   oryzae neutral protease II, a unique metalloprotease.";
RL   Mol. Gen. Genet. 228:97-103(1991).
CC   -!- FUNCTION: THERMOSTABLE METALLOPROTEASE. SHOWS HIGH ACTIVITIES ON
CC       BASIC NUCLEAR SUBSTRATES SUCH AS HISTONE AND PROTAMINE.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage of bonds with
CC       hydrophobic residues in P1'; also 3-Asn-|-Gln-4 and 8-Gln-|-Ser-9
CC       bonds in insulin B chain.
CC   -!- COFACTOR: BINDS 1 ZINC ION.
CC   -!- PTM: PROBABLY POSSESSES THREE DISULFIDE BONDS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M35 (ZINC
CC       METALLOPROTEASE); ALSO KNOWN AS THE DEUTEROLYSIN SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; S53810; AAB19701.1; -.
DR   MEROPS; M35.002; -.
DR   InterPro; IPR001384; Peptidase_M35.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF02102; Peptidase_M35; 1.
DR   PRINTS; PR00768; DEUTEROLYSIN.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Zymogen.
FT   SIGNAL       1     19       POTENTIAL.
FT   PROPEP      20    175
FT   CHAIN      176    352       NEUTRAL PROTEASE II.
FT   METAL      303    303       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE   304    304       BY SIMILARITY.
FT   METAL      307    307       ZINC (CATALYTIC) (BY SIMILARITY).
SQ   SEQUENCE   352 AA;  37517 MW;  070C5131335B7F44 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 352;
  Best Local Similarity   66.7%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    240 EAGSTS 245


RESULT   7
AR13_NEUCR
ID   AR13_NEUCR     STANDARD;      PRT;   363 AA.
AC   Q01356;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Amino acid transporter arg-13.
GN   ARG-13.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDA;
RX   MEDLINE=96400914; PubMed=8807290;
RA   Liu Q., Dunlap J.C.;
RT   "Isolation and analysis of the arg-13 gene of Neurospora crassa.";
RL   Genetics 143:1163-1174(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Liu Q., Luo X.;
RT   "Phenotypic rescue of Saccharomyces cerevisiae arg11 mutant by
RT   Neurospora crassa arg-13 cDNA.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR ARGININE BIOSYNTHESIS. MAY PARTICIPATE IN
CC       THE EXPORT OF MATRIX-MADE ORNITHINE INTO THE CYTOSOL (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       inner membrane (Potential).
CC   -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; L36378; AAC37500.1; -.
DR   EMBL; AF279268; AAF87777.1; -.
DR   InterPro; IPR001993; Mitoch_carrier.
DR   Pfam; PF00153; mito_carr; 3.
DR   PROSITE; PS00215; MITOCH_CARRIER; 2.
KW   Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT   TRANSMEM    141    161       POTENTIAL.
FT   TRANSMEM    266    283       POTENTIAL.
FT   TRANSMEM    334    353       POTENTIAL.
SQ   SEQUENCE   363 AA;  39401 MW;  8B87A937F6D37DC0 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 363;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     14 EAGAAS 19


RESULT   8
LEU3_BACME
ID   LEU3_BACME     STANDARD;      PRT;   367 AA.
AC   P41019;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE   (IMDH) (3-IPM-DH).
GN   LEUB OR LEUC.
OS   Bacillus megaterium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1404;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 319;
RX   MEDLINE=94288995; PubMed=7764969;
RA   Meinhardt F., Busskamp M., Wittchen K.D.;
RT   "Cloning and sequencing of the leu C and npr M genes and a putative
RT   spo IV gene from Bacillus megaterium DSM319.";
RL   Appl. Microbiol. Biotechnol. 41:344-351(1994).
CC   -!- CATALYTIC ACTIVITY: 3-CARBOXY-2-HYDROXY-4-METHYLPENTANOATE +
CC       NAD(+) = 3-CARBOXY-4-METHYL-2-OXOPENTANOATE + NADH (THE PRODUCT
CC       DECARBOXYLATES TO 4-METHYL-2-OXOPENTANOATE).
CC   -!- PATHWAY: THIRD STEP IN LEUCINE BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
CC       DEHYDROGENASES FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X65184; CAA46295.1; -.
DR   PIR; S38506; S38506.
DR   HSSP; P12010; 2AYQ.
DR   InterPro; IPR001804; Isodh.
DR   Pfam; PF00180; isodh; 1.
DR   PROSITE; PS00470; IDH_IMDH; 1.
KW   Oxidoreductase; Leucine biosynthesis; NAD.
SQ   SEQUENCE   367 AA;  39942 MW;  DC04D48E0EEAB0DD CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 367;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy      1 eagxxs 6
          |||  |
Db     48 EAGSSS 53


RESULT   9
CARP_SACFI
ID   CARP_SACFI     STANDARD;      PRT;   390 AA.
AC   P22929;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Acid protease precursor (EC 3.4.23.-).
GN   PEP1.
OS   Saccharomycopsis fibuligera (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX   NCBI_TaxID=4944;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hirata D., Fukui S., Yamashita I.;
RT   "Nucleotide sequence of the secretable acid protease gene PEPI in the
RT   yeast Saccharomycopsis fibuligera.";
RL   Agric. Biol. Chem. 52:2647-2649(1988).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC       EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00313; BAA00215.1; -.
DR   PIR; JT0334; JT0334.
DR   HSSP; P32329; 1YPS.
DR   InterPro; IPR001969; Asp_protease.
DR   InterPro; IPR001461; Pepsin.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    390       ACID PROTEASE.
FT   ACT_SITE     93     93       BY SIMILARITY.
FT   ACT_SITE    282    282       BY SIMILARITY.
SQ   SEQUENCE   390 AA;  41263 MW;  350BF97116C54796 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 390;
  Best Local Similarity   66.7%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    266 EAGSSS 271


RESULT  10
CAR4_DICDI
ID   CAR4_DICDI     STANDARD;      PRT;   443 AA.
AC   Q9TX43;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cyclic AMP receptor 4.
GN   CARD OR CAR4.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NC-4;
RX   MEDLINE=95047357; PubMed=7958880;
RA   Louis J.M., Ginsburg G.T., Kimmel A.R.;
RT   "The cAMP receptor CAR4 regulates axial patterning and cellular
RT   differentiation during late development of Dictyostelium.";
RL   Genes Dev. 8:2086-2096(1994).
CC   -!- FUNCTION: RECEPTOR FOR CAMP. REGULATES AXIAL PATTERNING AND
CC       CELLULAR DIFFERENTIATION DURING LATE DEVELOPMENT. THE ACTIVITY OF
CC       THIS RECEPTOR IS MEDIATED BY G PROTEINS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- DEVELOPMENTAL STAGE: INITIALLY EXPRESSED DURING TIP ELONGATION AND
CC       CONTINUES TO ACCUMULATE INTO CULMINATION.
CC   -!- PTM: CARBOXYL-TERMINAL SER OR THR RESIDUES MAY BE PHOSPHORYLATED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 5 OF G-PROTEIN COUPLED RECEPTORS.
DR   GCRDb; GCR_0277; -.
DR   DictyDb; DD0????; carD.
DR   InterPro; IPR000848; GPCR_cAMP.
DR   InterPro; IPR000832; GPCR_secretin.
DR   Pfam; PF00002; 7tm_2; 1.
DR   PRINTS; PR00247; GPCRCAMP.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Multigene family.
FT   DOMAIN        1     11       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     12     32       1 (POTENTIAL).
FT   DOMAIN       33     44       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     45     65       2 (POTENTIAL).
FT   DOMAIN       66     89       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     90    110       3 (POTENTIAL).
FT   DOMAIN      111    119       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    120    140       4 (POTENTIAL).
FT   DOMAIN      141    161       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    162    182       5 (POTENTIAL).
FT   DOMAIN      183    208       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    209    229       6 (POTENTIAL).
FT   DOMAIN      230    263       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    264    284       7 (POTENTIAL).
FT   DOMAIN      285    443       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      293    443       ASN-RICH.
FT   DOMAIN      293    298       POLY-ASN.
FT   DOMAIN      333    340       POLY-GLN.
FT   DOMAIN      343    353       POLY-GLN.
FT   DOMAIN      412    426       POLY-ASN.
SQ   SEQUENCE   443 AA;  51456 MW;  CDF3A9DEE5A5BBE2 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 443;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     72 EAGSTS 77


RESULT  11
VGLX_HSVBS
ID   VGLX_HSVBS     STANDARD;      PRT;   444 AA.
AC   Q08103;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein GX precursor.
OS   Bovine herpesvirus type 1.2 (strain ST).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=45407;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94167875; PubMed=8122370;
RA   Leung-Tack P., Audonnet J.F., Riviere M.;
RT   "The complete DNA sequence and the genetic organization of the short
```

```
RT   unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL   Virology 199:409-421(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z23068; CAA80603.1; -.
DR   InterPro; IPR003363; Herpes_gG.
DR   Pfam; PF02400; Herpes_gG; 1.
KW   Glycoprotein; Transmembrane; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    444       GLYCOPROTEIN GX.
FT   TRANSMEM    390    414       POTENTIAL.
FT   CARBOHYD    117    117       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    240    240       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    335    335       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   444 AA;  46708 MW;  0145942AA35B05CB CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 444;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  `|
Db    267 EAGSAS 272


RESULT  12
YU88_MYCTU
ID   YU88_MYCTU     STANDARD;      PRT;   474 AA.
AC   O53305;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 50.9 kDa protein Rv3088.
GN   RV3088 OR MT3173 OR MTV013.09.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL021309; CAA16146.1; -.
DR   EMBL; AE007134; AAK47509.1; -.
DR   TIGR; MT3173; -.
DR   TubercuList; Rv3088; -.
DR   InterPro; IPR004255; UPF0089.
DR   Pfam; PF03007; UPF0089; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   474 AA;  50886 MW;  36832D972BE3851A CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 474;
  Best Local Similarity   66.7%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    460 EAGTTS 465


RESULT  13
NODT_RHILV
ID   NODT_RHILV     STANDARD;      PRT;   482 AA.
AC   P15727;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nodulation protein T precursor.
GN   NODT.
OS   Rhizobium leguminosarum (biovar viciae).
OG   Plasmid sym pRL1JI.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=387;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8401;
RX   MEDLINE=90151607; PubMed=2303029;
RA   Economou A., Hamilton W.D.O., Johnston A.W.B., Downie J.A.;
RT   "The Rhizobium nodulation gene nodO encodes a Ca2(+)-binding protein
RT   that is exported without N-terminal cleavage and is homologous to
RT   haemolysin and related proteins.";
RL   EMBO J. 9:349-354(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90251164; PubMed=2338917;
RA   Surin B.P., Watson J.M., Hamilton W.D.O., Economou A., Downie J.A.;
RT   "Molecular characterization of the nodulation gene, nodT, from two
RT   biovars of Rhizobium leguminosarum.";
RL   Mol. Microbiol. 4:245-252(1990).
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC       (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FUSA/NODT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17285; CAA35177.1; -.
DR   PIR; S08384; S08384.
DR   PIR; S10133; S10133.
```

```
DR   InterPro; IPR003423; OEP.
DR   Pfam; PF02321; OEP; 2.
KW   Plasmid; Nodulation; Lipoprotein; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    482       NODULATION PROTEIN T.
FT   LIPID        18     18       N-ACYL DIGLYCERIDE (POTENTIAL).
SQ   SEQUENCE   482 AA;  51488 MW;  D555B46566E54F82 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 482;
  Best Local Similarity   66.7%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    201 EAGAAS 206


RESULT  14
ATF2_RAT
ID   ATF2_RAT       STANDARD;      PRT;   487 AA.
AC   Q00969; Q62870;
DT   01-APR-1993 (Rel. 25, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cyclic-AMP-dependent transcription factor ATF-2 (Activating
DE   transcription factor 2) (cAMP response element binding protein CRE-
DE   BP1).
GN   ATF2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RA   Muramatsu S.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RX   MEDLINE=91332085; PubMed=1714459;
RA   Kageyama R., Sasai Y., Nakanishi S.;
RT   "Molecular characterization of transcription factors that bind to the
RT   cAMP responsive region of the substance P precursor gene. cDNA
RT   cloning of a novel C/EBP-related factor.";
RL   J. Biol. Chem. 266:15525-15531(1991).
CC   -!- FUNCTION: A TRANSCRIPTIONAL ACTIVATOR, PROBABLY CONSTITUTIVE,
CC       WHICH BINDS TO THE CAMP RESPONSIVE REGION (CRE) (CONSENSUS:
CC       5'GTGACGT(A/C)(A/G)-3'); A SEQUENCE PRESENT IN MANY VIRAL AND
CC       CELLULAR PROMOTERS.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U38938; AAA93263.1; -.
DR   EMBL; M65148; AAA42013.1; -.
DR   PIR; A39429; A39429.
DR   HSSP; P08047; 1SP2.
DR   TRANSFAC; T01382; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF00170; bZIP; 1.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00338; BRLZ; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00036; BZIP_BASIC; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   Transcription regulation; DNA-binding; Activator; Phosphorylation;
KW   Nuclear protein; Alternative splicing; Zinc-finger; Metal-binding.
FT   ZN_FING       7     31       C2H2-TYPE.
FT   DNA_BIND    333    356       BASIC MOTIF.
FT   DOMAIN      362    390       LEUCINE-ZIPPER.
FT   MOD_RES      51     51       PHOSPHORYLATION (BY MAPK14) (BY
FT                                SIMILARITY).
FT   MOD_RES      53     53       PHOSPHORYLATION (BY MAPK14) (BY
FT                                SIMILARITY).
FT   VARSPLIC    132    229       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   487 AA;  52286 MW;  4ED95B106DF5F9EE CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 487;
  Best Local Similarity   66.7%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    452 EAGATS 457


RESULT  15
MERA_ACICA
ID   MERA_ACICA     STANDARD;      PRT;   561 AA.
AC   Q52109;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Mercuric reductase (EC 1.16.1.1) (Hg(II) reductase).
GN   MERA.
OS   Acinetobacter calcoaceticus.
OG   Plasmid pKLH2.
OC   Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC   Acinetobacter.
OX   NCBI_TaxID=471;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94134837; PubMed=8302940;
RA   Kholodii G.Y., Lomovskaya O.L., Gorlenko Z.M., Mindlin S.Z.,
RA   Yurieva O.V., Nikiforov V.G.;
RT   "Molecular characterization of an aberrant mercury resistance
RT   transposable element from an environmental Acinetobacter strain.";
RL   Plasmid 30:303-308(1993).
CC   -!- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC       BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE. MERA
CC       PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC   -!- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
CC   -!- COFACTOR: FAD.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC       OXIDOREDUCTASES CLASS-I.
CC   -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF213017; AAA19682.1; -.
DR   HSSP; Q04656; 1AW0.
DR   InterPro; IPR001327; FAD_pyr_redox.
DR   InterPro; IPR001934; HMA.
```

```
DR   InterPro; IPR000815; Hg_reductase.
DR   InterPro; IPR001100; pyr_redox.
DR   InterPro; IPR004099; pyr_redox_dim.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00070; pyr_redox; 1.
DR   Pfam; PF02852; pyr_redox_dim; 1.
DR   PRINTS; PR00368; FADPNR.
DR   PRINTS; PR00945; HGRDTASE.
DR   PRINTS; PR00411; PNDRDTASEI.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
DR   PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW   Mercuric resistance; Oxidoreductase; Flavoprotein; FAD; NADP;
KW   Mercury; Redox-active center; Metal-binding; Plasmid.
FT   DOMAIN        1     66       HMA.
FT   NP_BIND     100    130       FAD (ADP PART) (PROBABLE).
FT   DISULFID    136    141       REDOX-ACTIVE.
FT   NP_BIND     393    403       FAD (FLAVIN PART) (BY SIMILARITY).
FT   METAL       558    558       HG(2+) (POTENTIAL).
FT   METAL       559    559       HG(2+) (POTENTIAL).
SQ   SEQUENCE   561 AA;  58558 MW;  111E02A702C157D6 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 561;
  Best Local Similarity   66.7%;  Pred. No. 4.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     45 EAGTSS 50


RESULT  16
MERA_ENTAG
ID   MERA_ENTAG     STANDARD;      PRT;   561 AA.
AC   P94702;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Mercuric reductase (EC 1.16.1.1) (Hg(II) reductase).
GN   MERA.
OS   Enterobacter agglomerans (Pantoea agglomerans).
OG   Plasmid pKLH272.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Pantoea.
OX   NCBI_TaxID=549;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97303088; PubMed=9159519;
RA   Yurieva O., Kholodii G., Minakhin L., Gorlenko Z., Kalyaeva E.,
RA   Mindlin S., Nikiforov V.;
RT   "Intercontinental spread of promiscuous mercury-resistance
RT   transposons in environmental bacteria.";
RL   Mol. Microbiol. 24:321-329(1997).
CC   -!- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC       BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE. MERA
CC       PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC   -!- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
CC   -!- COFACTOR: FAD.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC       OXIDOREDUCTASES CLASS-I.
CC   -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; Y08992; CAA70184.1; -.
DR   HSSP; Q04656; 1AW0.
DR   InterPro; IPR001327; FAD_pyr_redox.
DR   InterPro; IPR001934; HMA.
DR   InterPro; IPR000815; Hg_reductase.
DR   InterPro; IPR000205; NAD_binding.
DR   InterPro; IPR001100; pyr_redox.
DR   InterPro; IPR004099; pyr_redox_dim.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00070; pyr_redox; 1.
DR   Pfam; PF02852; pyr_redox_dim; 1.
DR   PRINTS; PR00368; FADPNR.
DR   PRINTS; PR00945; HGRDTASE.
DR   PRINTS; PR00411; PNDRDTASEI.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
DR   PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW   Mercuric resistance; Oxidoreductase; Flavoprotein; FAD; NADP;
KW   Mercury; Redox-active center; Metal-binding; Plasmid.
FT   DOMAIN        1     66       HMA.
FT   NP_BIND     100    130       FAD (ADP PART) (PROBABLE).
FT   DISULFID    136    141       REDOX-ACTIVE.
FT   NP_BIND     393    403       FAD (FLAVIN PART) (BY SIMILARITY).
FT   METAL       558    558       HG(2+) (POTENTIAL).
FT   METAL       559    559       HG(2+) (POTENTIAL).
SQ   SEQUENCE   561 AA;  58785 MW;  FABA07D7EC2F13C8 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 561;
  Best Local Similarity   66.7%;  Pred. No. 4.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     45 EAGTSS 50


RESULT  17
AC22_STRCO
ID   AC22_STRCO     STANDARD;      PRT;   578 AA.
AC   P46105;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable actinorhodin transporter.
GN   ACTII-2 OR SCBAC28G1.09.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91347376; PubMed=1878971;
RA   Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;
RT   "The act cluster contains regulatory and antibiotic export genes,
RT   direct targets for translational control by the bldA tRNA gene of
RT   Streptomyces.";
RL   Cell 66:769-780(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Warren T., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,
RA   Rajandream M.A.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROMOTES THE EFFLUX OF ACTINORHODIN.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC       AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M64683; AAA26690.1; -.
DR   EMBL; AL593842; CAC44196.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
KW   Antibiotic resistance; Transport; Transmembrane.
FT   TRANSMEM     78     98       POTENTIAL.
FT   TRANSMEM    109    129       POTENTIAL.
FT   TRANSMEM    135    155       POTENTIAL.
FT   TRANSMEM    170    190       POTENTIAL.
FT   TRANSMEM    202    222       POTENTIAL.
FT   TRANSMEM    232    252       POTENTIAL.
FT   TRANSMEM    259    279       POTENTIAL.
FT   TRANSMEM    306    326       POTENTIAL.
FT   TRANSMEM    341    361       POTENTIAL.
FT   TRANSMEM    369    389       POTENTIAL.
FT   TRANSMEM    444    464       POTENTIAL.
FT   TRANSMEM    546    566       POTENTIAL.
SQ   SEQUENCE   578 AA;  59772 MW;  E6C1DC75E6038B92 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 578;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    430 EAGTAS 435


RESULT  18
DNK2_SYNY3
ID   DNK2_SYNY3     STANDARD;      PRT;   636 AA.
AC   P22358;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chaperone protein dnaK2 (Heat shock protein 70-2) (Heat shock 70 kDa
DE   protein 2) (HSP70-2).
GN   DNAK2 OR DNAK OR SLL0170.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91093185; PubMed=1670771;
RA   Chitnis P.R., Nelson N.;
RT   "Molecular cloning of the genes encoding two chaperone proteins of
RT   the cyanobacterium Synechocystis sp. PCC 6803.";
RL   J. Biol. Chem. 266:58-65(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC   -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK.
CC   -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M57518; AAA27287.1; -.
DR   EMBL; D63999; BAA10059.1; -.
DR   PIR; C39025; C39025.
DR   HSSP; P04475; 1DG4.
DR   InterPro; IPR001023; HSP70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   PROSITE; PS00297; HSP70_1; 1.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   Chaperone; ATP-binding; Heat shock; Multigene family;
KW   Complete proteome.
SQ   SEQUENCE   636 AA;  67614 MW;  33AE4CECBA28F40A CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 636;
  Best Local Similarity   66.7%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    615 EAGTSS 620


RESULT  19
TOXA_PSEAE
ID   TOXA_PSEAE     STANDARD;      PRT;   638 AA.
AC   P11439; Q9I4I7;
DT   01-OCT-1989 (Rel. 12, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Exotoxin A precursor (NAD-dependent ADP-ribosyltransferase
DE   (EC 2.4.2.-)).
GN   ETA OR PA1148.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 26-53.
RX   MEDLINE=84194063; PubMed=6201861;
RA   Gray G.L., Smith D.H., Baldridge J.S., Harkins R.N., Vasil M.L.,
RA   Chen E.Y., Heyneker H.L.;
RT   "Cloning, nucleotide sequence, and expression in Escherichia coli of
RT   the exotoxin A structural gene of Pseudomonas aeruginosa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:2645-2649(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
RN   [3]
RP   ACTIVE SITE.
RX   MEDLINE=87250491; PubMed=2885323;
RA   Carroll S.F., Collier R.J.;
RT   "Active site of Pseudomonas aeruginosa exotoxin A. Glutamic acid 553
RT   is photolabeled by NAD and shows functional homology with glutamic
RT   acid 148 of diphtheria toxin.";
RL   J. Biol. Chem. 262:8707-8711(1987).
RN   [4]
RP   DOMAINS.
RX   MEDLINE=90375493; PubMed=2118903;
RA   Chaudhary V.K., Jinno Y., Galo M.G., Fitzgerald D., Pastan I.;
RT   "Mutagenesis of Pseudomonas exotoxin in identification of sequences
```

```
RT   responsible for the animal toxicity.";
RL   J. Biol. Chem. 265:16306-16310(1990).
RN   [5]
RP   DOMAINS.
RX   MEDLINE=91006124; PubMed=2170123;
RA   Bourdenet S., Vacheron M.-J., Guinand M., Michel G., Arminjon F.;
RT   "Biochemical and immunochemical studies of proteolytic fragments of
RT   exotoxin A from Pseudomonas aeruginosa.";
RL   Eur. J. Biochem. 192:379-385(1990).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 424-638.
RX   MEDLINE=96016159; PubMed=7568123;
RA   Li M., Dyda F., Benhar I., Pastan I., Davies D.R.;
RT   "The crystal structure of Pseudomonas aeruginosa exotoxin domain III
RT   with nicotinamide and AMP: conformational differences with the intact
RT   exotoxin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9308-9312(1995).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 424-638.
RX   MEDLINE=96293446; PubMed=8692916;
RA   Li M., Dyda F., Benhar I., Pastan I., Davies D.R.;
RT   "Crystal structure of the catalytic domain of Pseudomonas exotoxin A
RT   complexed with a nicotinamide adenine dinucleotide analog:
RT   implications for the activation process and for ADP ribosylation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:6902-6906(1996).
CC   -!- FUNCTION: THIS TOXIN IS AN NAD-DEPENDENT ADP-RIBOSYLTRANSFERASE.
CC       IT CATALYZES THE TRANSFER OF THE ADP RIBOSYL MOIETY OF OXIDIZED
CC       NAD ONTO ELONGATION FACTOR 2 (EF-2) THUS ARRESTING PROTEIN
CC       SYNTHESIS.
CC   -!- PTM: THE 8 CYSTEINES PARTICIPATE IN INTRACHAIN DISULFIDE BONDS.
CC   -!- SIMILARITY: REGIONAL SEQUENCE SIMILARITY AT THE ACTIVE SITE
CC       WITH DIPHTHERIA TOXIN (DT).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K01397; AAB59097.1; -.
DR   EMBL; AE004544; AAG04537.1; -.
DR   PIR; A30347; A30347.
DR   PDB; 1AER; 10-JUN-96.
DR   PDB; 1DMA; 15-SEP-95.
KW   Toxin; Signal; Transferase; Glycosyltransferase; NAD; 3D-structure;
KW   Complete proteome.
FT   SIGNAL       1     25
FT   CHAIN       26    638       EXOTOXIN A.
FT   DOMAIN      26    277       IA (REQUIRED FOR TARGET CELL
FT                               RECOGNITION).
FT   DOMAIN     278    389       II (REQUIRED FOR TRANSLOCATION IN TARGET
FT                               CELL CYTOPLASM).
FT   DOMAIN     390    429       IB.
FT   DOMAIN     430    638       III (REQUIRED FOR ADP-RIBSOSYL ACTIVITY).
FT   ACT_SITE   465    465
FT   ACT_SITE   578    578       INTERACT WITH NAD.
FT   CONFLICT     4      4       T -> I (IN REF. 1).
FT   CONFLICT    22     22       F -> S (IN REF. 1).
FT   CONFLICT   204    204       A -> T (IN REF. 1).
FT   CONFLICT   389    389       S -> N (IN REF. 1).
FT   CONFLICT   432    432       I -> V (IN REF. 1).
FT   CONFLICT   540    540       G -> S (IN REF. 1).
SQ   SEQUENCE   638 AA;  69284 MW;  7B9AAD56A27C700A CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 638;
  Best Local Similarity   66.7%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    384 EAGAAS 389


RESULT  20
SGT1_ARATH
ID   SGT1_ARATH    . STANDARD;      PRT;   643 AA.
AC   Q9LSM5;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   SGT1 protein homolog At5g65490.
GN   AT5G65490 OR K19O4.2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE SGT1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB026638; BAA98169.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   643 AA;  73161 MW;  9F9DD990F65B4C4F CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 643;
  Best Local Similarity   66.7%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    445 EAGSSS 450


RESULT  21
EOMD_XENLA
ID   EOMD_XENLA     STANDARD;      PRT;   692 AA.
AC   P79944;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Eomesodermin.
GN   EOMES.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Gastrula;
RX   MEDLINE=97133207; PubMed=8978604;
RA   Ryan K., Garrett N., Mitchell A., Gurdon J.B.;
RT   "Eomesodermin, a key early gene in Xenopus mesoderm differentiation.";
RL   Cell 87:989-1000(1996).
CC   -!- FUNCTION: INVOLVED IN MESODERM DIFFERENTIATION. ACTIVATES WNT8,
CC       BRACHYURY, CHD AND MIX.1 EXPRESSION.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT OR JUST AFTER MIDBLASTULA
```

```
CC       TRANSITION (STAGE 8). MAXIMALLY EXPRESSED AT STAGE 10 AS AN
CC       EQUITORIAL MESODERM BAND, MORE PROMINENTLY ON THE DORSAL SIDE
CC       AND AROUND THE INVAGINATING DORSAL LIP.
CC   -!- INDUCTION: BY ACTIVIN.
CC   -!- DOMAIN: CONTAINS 13 S-P-X-X REPEATS.
CC   -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U75996; AAC60061.1; -.
DR   HSSP; P24781; 1XBR.
DR   InterPro; IPR001699; T-box.
DR   Pfam; PF00907; T-box; 1.
DR   PRINTS; PR00937; TBOX.
DR   SMART; SM00425; TBOX; 1.
DR   PROSITE; PS01283; TBOX_1; 1.
DR   PROSITE; PS01264; TBOX_2; 1.
DR   PROSITE; PS50252; TBOX_3; 1.
KW   Developmental protein; Transcription regulation; DNA-binding;
KW   Nuclear protein; Repeat.
FT   DNA_BIND    263    443       T-BOX.
SQ   SEQUENCE   692 AA;  75943 MW;  9D129A67F6357989 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 692;
  Best Local Similarity   66.7%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     90 EAGSSS 95


RESULT  22
PGCB_MOUSE
ID   PGCB_MOUSE     STANDARD;      PRT;   883 AA.
AC   Q61361;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Brevican core protein precursor.
GN   BCAN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Brain;
RX   MEDLINE=97432816; PubMed=9286696;
RA   Rauch U., Meyer H., Brakebusch C., Seidenbecher C., Gundelfinger E.D.,
RA   Beier D.R., Fassler R.;
RT   "Sequence and chromosomal localization of the mouse brevican gene.";
RL   Genomics 44:15-21(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC       THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC       STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCANS.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
CC       SIMILARITY).
CC   -!- TISSUE SPECIFICITY: BRAIN (BY SIMILARITY).
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X87096; CAA60575.1; -.
DR   HSSP; P20693; 1HLJ.
DR   MGD; MGI:1096385; Bcan.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   InterPro; IPR001304; lectin_c.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW   EGF-like domain; Repeat; Immunoglobulin domain.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23    883       BREVICAN CORE PROTEIN.
FT   DOMAIN      32    157       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN     173    250       LINK 1.
FT   DOMAIN     271    352       LINK 2.
FT   DOMAIN     622    658       EGF-LIKE.
FT   DOMAIN     658    786       C-TYPE LECTIN.
FT   DOMAIN     787    851       SUSHI.
FT   DISULFID    56    136       BY SIMILARITY.
FT   DISULFID   178    249       BY SIMILARITY.
FT   DISULFID   202    223       BY SIMILARITY.
FT   DISULFID   276    351       BY SIMILARITY.
FT   DISULFID   300    321       BY SIMILARITY.
FT   DISULFID   626    637       BY SIMILARITY.
FT   DISULFID   631    646       BY SIMILARITY.
FT   DISULFID   648    657       BY SIMILARITY.
FT   DISULFID   664    675       BY SIMILARITY.
FT   DISULFID   692    784       BY SIMILARITY.
FT   DISULFID   760    776       BY SIMILARITY.
FT   DISULFID   791    834       BY SIMILARITY.
FT   DISULFID   820    847       BY SIMILARITY.
FT   CARBOHYD   129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   883 AA;  96013 MW;  CC2C33C97B453E45 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 883;
  Best Local Similarity   66.7%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    558 EAGSSS 563
```

```
RESULT  23
PGCB_RAT
ID   PGCB_RAT       STANDARD;      PRT;   883 AA.
AC   P55068; Q63040; Q62860; Q63513;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Brevican core protein precursor (Brain enriched hyaluronan binding
DE   protein) (BEHAB protein).
GN   BCAN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX   MEDLINE=96070828; PubMed=7592978;
RA   Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C.,
RA   Gundelfinger E.D.;
RT   "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as
RT   secreted and cell surface glycosylphosphatidylinositol-anchored
RT   isoforms.";
RL   J. Biol. Chem. 270:27206-27212(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RC   TISSUE=Brain;
RX   MEDLINE=96074575; PubMed=7488217;
RA   Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamaguchi Y.;
RT   "cDNA cloning and the identification of an aggrecanase-like cleavage
RT   site in rat brevican.";
RL   Biochem. Biophys. Res. Commun. 216:957-963(1995).
RN   [3]
RP   SEQUENCE OF 1-423 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX   MEDLINE=94216386; PubMed=7512973;
RA   Jaworski D.M., Kelly G.M., Hockfield S.;
RT   "BEHAB, a new member of the proteoglycan tandem repeat family of
RT   hyaluronan-binding proteins that is restricted to the brain.";
RL   J. Cell Biol. 125:495-509(1994).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC       THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC       STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCANS. THE
CC       GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-
CC       BEARING CELL SURFACE RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM
CC       ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SECRETED FORM (SHOWN HERE) AND
CC       A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.
CC       GPI-ANCHORED ISOFORM INCREASES AFTER DAY P8.
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364
CC       ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79881; CAA56255.1; -.
DR   EMBL; X86406; CAA60160.1; -.
DR   EMBL; U37142; AAA87847.1; -.
DR   EMBL; Z28366; CAA82215.1; ALT_FRAME.
DR   HSSP; P20693; 1HLJ.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   InterPro; IPR001304; lectin_c.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW   EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;
KW   GPI-anchor.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    883       BREVICAN CORE PROTEIN.
FT   DOMAIN       32    157       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      173    250       LINK 1.
FT   DOMAIN      271    352       LINK 2.
FT   DOMAIN      622    658       EGF-LIKE.
FT   DOMAIN      658    786       C-TYPE LECTIN.
FT   DOMAIN      787    851       SUSHI.
FT   DISULFID     56    136       BY SIMILARITY.
FT   DISULFID    178    249       BY SIMILARITY.
FT   DISULFID    202    223       BY SIMILARITY.
FT   DISULFID    276    351       BY SIMILARITY.
FT   DISULFID    300    321       BY SIMILARITY.
FT   DISULFID    626    637       BY SIMILARITY.
FT   DISULFID    631    646       BY SIMILARITY.
FT   DISULFID    648    657       BY SIMILARITY.
FT   DISULFID    791    834       BY SIMILARITY.
FT   DISULFID    820    847       BY SIMILARITY.
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    625    645       DCIPSPCHNGGTCLEEKEGFR -> NSAEGSMPAFLLFLLL
FT                                QLWDT (IN GPI-ANCHORED ISOFORM).
FT   VARSPLIC    646    883       MISSING (IN GPI-ANCHORED ISOFORM).
FT   CONFLICT     51     52       AL -> WV (IN REF. 3).
FT   CONFLICT    503    503       V -> L (IN REF. 2).
FT   CONFLICT    518    519       TV -> PA (IN REF. 2).
FT   CONFLICT    526    526       G -> R (IN REF. 2).
FT   CONFLICT    541    541       G -> A (IN REF. 2).
FT   CONFLICT    556    556       R -> S (IN REF. 2).
FT   CONFLICT    573    573       E -> A (IN REF. 2).
FT   CONFLICT    583    583       V -> L (IN REF. 2).
FT   CONFLICT    649    649       V -> L (IN REF. 2).
FT   CONFLICT    670    670       P -> A (IN REF. 2).
FT   CONFLICT    738    738       P -> A (IN REF. 2).
FT   CONFLICT    809    809       R -> A (IN REF. 2).
SQ   SEQUENCE   883 AA;  96057 MW;  AC7ACC40CB53ED37 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 883;
  Best Local Similarity   66.7%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
```

```
Db    558 EAGSSS 563


RESULT  24
SRCA_RABIT
ID   SRCA_RABIT     STANDARD;      PRT;   908 AA.
AC   P13666;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sarcalumenin precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89345602; PubMed=2762314;
RA   Leberer E., Charuk J.H.M., Green N.M., Maclennan D.H.;
RT   "Molecular cloning and expression of cDNA encoding a lumenal calcium
RT   binding glycoprotein from sarcoplasmic reticulum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:6047-6051(1989).
RN   [2]
RP   SEQUENCE OF 1-19 AND 458-908 FROM N.A.
RX   MEDLINE=89123480; PubMed=2521635;
RA   Leberer E., Charuk J.H.M., Clarke D.M., Green N.M., Zubrycka-Gaarn E.,
RA   McLennan D.H.;
RT   "Molecular cloning and expression of cDNA encoding the 53,000-dalton
RT   glycoprotein of rabbit skeletal muscle sarcoplasmic reticulum.";
RL   J. Biol. Chem. 264:3484-3493(1989).
CC   -!- FUNCTION: PERHAPS INVOLVED IN THE REGULATION OF CALCIUM TRANSPORT.
CC   -!- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM LUMEN. ASSOCIATED
CC       THROUGH CA(2+) WITH THE MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A CALCIUM-BINDING
CC       GLYCOPROTEIN/160 KDA (SHOWN HERE) AND A SECOND GLYCOPROTEIN/53
CC       KDA; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M25750; AAA31189.1; -.
DR   EMBL; J04480; AAA60730.1; -.
DR   PIR; A33280; A33280.
DR   PIR; A33312; A33312.
KW   Calcium-binding; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     19
FT   CHAIN        20    908       160 KDA SARCALUMENIN.
FT   CHAIN       458    908       53 KDA SARCALUMENIN.
FT   DOMAIN       20    457       ACIDIC DOMAIN, PROBABLY BINDS CALCIUM.
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    716    716       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    824    824       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     21    456       MISSING (IN 53 KDA ISOFORM).
FT   CONFLICT    474    474       Q -> E (IN REF. 2).
SQ   SEQUENCE   908 AA;  97920 MW;  A48CAA221AE1418B CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 908;
  Best Local Similarity   66.7%;  Pred. No. 7.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    397 EAGAAS 402


RESULT  25
POLS_IBDVA
ID   POLS_IBDVA     STANDARD;      PRT;  1012 AA.
AC   P08364;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Structural polyprotein [Contains: Major structural protein VP2;
DE   Nonstructural protein VP4; Minor structural protein VP3].
OS   Avian infectious bursal disease virus (strain Australian 002-73)
OS   (IBDV).
OC   Viruses; dsRNA viruses; Birnaviridae; Avibirnavirus.
OX   NCBI_TaxID=10997;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86259073; PubMed=3014441;
RA   Hudson P.J., McKern N.M., Power B.E., Azad A.A.;
RT   "Genomic structure of the large RNA segment of infectious bursal
RT   disease virus.";
RL   Nucleic Acids Res. 14:5001-5012(1986).
RN   [2]
RP   SEQUENCE OF 703-1012 FROM N.A.
RX   MEDLINE=86220784; PubMed=3011501;
RA   Hudson P.J., McKern N.M., Fahey K.J., Azad A.A.;
RT   "Predicted sequence of the host-protective immunogen of infectious
RT   bursal disease virus.";
RL   FEBS Lett. 201:143-146(1986).
CC   -!- FUNCTION: SEGMENT A ENCODES A POLYPROTEIN, THAT IS PROCESSED INTO
CC       THE MAJOR STRUCTURAL PROTEINS OF THE VIRION VP2 AND VP3, AND INTO
CC       THE PUTATIVE PROTEASE VP4.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X03993; CAA27629.1; ALT_INIT.
DR   PIR; A24382; GNXSAU.
DR   MEROPS; S50.002; -.
DR   InterPro; IPR002662; Birna_VP2.
DR   InterPro; IPR002663; Birna_VP3.
DR   InterPro; IPR002664; Birna_VP4.
DR   Pfam; PF01766; Birna_VP2; 1.
DR   Pfam; PF01767; Birna_VP3; 1.
DR   Pfam; PF01768; Birna_VP4; 1.
KW   Polyprotein; Structural protein; Nonstructural protein; Hydrolase;
KW   Protease.
FT   CHAIN         1    452       MAJOR STRUCTURAL PROTEIN VP2.
FT   CHAIN       454    722       NONSTRUCTURAL PROTEIN VP4 (PROTEASE).
FT   CHAIN       724   1012       MINOR STRUCTURAL PROTEIN VP3.
SQ   SEQUENCE   1012 AA;  109503 MW;  D9320A90459DE8F6 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1012;
  Best Local Similarity   66.7%;  Pred. No. 8.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    484 EAGAAS 489


RESULT  26
TSCC_HUMAN
ID   TSCC_HUMAN     STANDARD;      PRT;  1021 AA.
AC   P55017;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thiazide-sensitive sodium-chloride cotransporter (NA-CL symporter).
GN   SLC12A3 OR TSC.
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS GS.
RX   MEDLINE=96122035; PubMed=8528245;
RA   Simon D.B., Nelson-Williams C., Bia M.J., Ellison D., Karet F.E.,
RA   Molina A.M., Vaara I., Iwata F., Cushner H.M., Koolen M., Gainza F.J.,
RA   Gitelman H.J., Lifton R.P.;
RT   "Gitelman's variant of Bartter's syndrome, inherited hypokalaemic
RT   alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl
RT   cotransporter.";
RL   Nat. Genet. 12:24-30(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=97001149; PubMed=8812482;
RA   Mastroianni N., de Fusco M., Zollo M., Arrigo G., Zuffardi O.,
RA   Bettinelli A., Ballabio A., Casari G.;
RT   "Molecular cloning, expression pattern, and chromosomal localization
RT   of the human Na-Cl thiazide-sensitive cotransporter (SLC12A3).";
RL   Genomics 35:486-493(1996).
CC   -!- FUNCTION: ELECTRICALLY SILENT TRANSPORTER SYSTEM WHICH IS A
CC       MEDIATOR OF SODIUM AND CHLORIDE REABSORPTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PREDOMINANT IN KIDNEY.
CC   -!- DISEASE: DEFECTS IN SLC12A3 ARE THE CAUSE OF GITELMAN'S SYNDROME
CC       (GS), AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY DIVERSE
CC       ABNORMALITIES IN ELECTROLYTE HOMEOSTASIS INCLUDING HYPOKALAEMIC
CC       METABOLIC ALKALOSIS. GS IS A SUBSET OF BARTTER'S SYNDROME.
CC   -!- SIMILARITY: BELONGS TO THE SLC12A FAMILY OF TRANSPORTERS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U44128; AAC50355.1; -.
DR   EMBL; X91220; CAA62613.1; -.
DR   MIM; 600968; -.
DR   MIM; 263800; -.
DR   InterPro; IPR002293; AA_rel_permease_1.
DR   InterPro; IPR002948; NaCl_transporter.
DR   PRINTS; PR01230; NACLTRNSPORT.
KW   Transport; Transmembrane; Glycoprotein; Disease mutation.
FT   DOMAIN        1    135       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    136    156       POTENTIAL.
FT   DOMAIN      157    158       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    159    179       POTENTIAL.
FT   DOMAIN      180    218       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    219    239       POTENTIAL.
FT   DOMAIN      240    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    262    282       POTENTIAL.
FT   DOMAIN      283    286       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    287    307       POTENTIAL.
FT   DOMAIN      308    339       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    340    360       POTENTIAL.
FT   DOMAIN      361    377       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    378    398       POTENTIAL.
FT   DOMAIN      399    452       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    453    473       POTENTIAL.
FT   DOMAIN      474    511       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    512    532       POTENTIAL.
FT   DOMAIN      533    534       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    535    555       POTENTIAL.
FT   DOMAIN      556    577       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    578    598       POTENTIAL.
FT   DOMAIN      599    660       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    661    681       POTENTIAL.
FT   DOMAIN      682   1021       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    426    426       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     209    209       R -> W (IN GS).
FT                                /FTId=VAR_007113.
FT   VARIANT     349    349       P -> L (IN GS).
FT                                /FTId=VAR_007114.
FT   VARIANT     421    421       C -> R (IN GS).
FT                                /FTId=VAR_007115.
FT   VARIANT     486    486       D -> N (IN GS).
FT                                /FTId=VAR_007116.
FT   VARIANT     496    496       G -> C (IN GS).
FT                                /FTId=VAR_007117.
FT   VARIANT     561    561       MISSING (IN GS).
FT                                /FTId=VAR_007118.
FT   VARIANT     588    588       A -> V (IN GS).
FT                                /FTId=VAR_007119.
FT   VARIANT     630    630       G -> V (IN GS).
FT                                /FTId=VAR_007120.
FT   VARIANT     655    655       R -> H (IN GS).
FT                                /FTId=VAR_007121.
FT   VARIANT     655    655       R -> L (IN GS).
FT                                /FTId=VAR_007122.
FT   VARIANT     728    728       A -> T (IN GS).
FT                                /FTId=VAR_007123.
FT   VARIANT     741    741       G -> R (IN GS).
FT                                /FTId=VAR_007124.
FT   VARIANT     850    850       L -> P (IN GS).
FT                                /FTId=VAR_007125.
FT   VARIANT     955    955       R -> Q (IN GS).
FT                                /FTId=VAR_007126.
FT   CONFLICT    459    460       AG -> VV (IN REF. 2).
FT   CONFLICT    766    766       E -> D (IN REF. 2).
FT   CONFLICT    807    807       V -> GARPSVSGAL (IN REF. 1).
SQ   SEQUENCE   1021 AA;  113138 MW;  D7ECE53DA6233821 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1021;
  Best Local Similarity   66.7%;  Pred. No. 8.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     121 EAGTSS 126


RESULT  27
ISWI_DROME
ID   ISWI_DROME     STANDARD;      PRT;  1027 AA.
AC   Q24368; Q9V6E8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Iswi protein (Imitation swi protein) (Nucleosome remodeling factor 140
DE   kDa subunit) (NURF-140) (CHRAC 140 kDa subunit).
GN   ISWI OR CG8625.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ISO-1;
RX   MEDLINE=94187693; PubMed=7908117;
RA   Elfring L.K., Deuring R., McCallum C.M., Peterson C.L., Tamkun J.W.;
RT   "Identification and characterization of Drosophila relatives of the
RT   yeast transcriptional activator SNF2/SWI2.";
RL   Mol. Cell. Biol. 14:2225-2234(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;*
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
```

```
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: COMPONENT OF THE NUCLEOSOME REMODELING FACTOR COMPLEX
CC       (NURF), A PROTEIN COMPLEX THAT FACILITATES THE PERTURBATION OF
CC       CHROMATIN STRUCTURE IN VITRO IN AN ATP-DEPENDENT MANNER. THE
CC       HYDROLYSIS OF ATP DURING THE REMODELING OF CHROMATIN IS LIKELY TO
CC       BE MEDIATED BY ISWI, RELEASING INORGANIC PHOSPHATE. IT IS ALSO A
CC       COMPONENT OF THE ATP-UTILIZING CHROMATIN ASSEMBLY AND REMODELING
CC       FACTOR (ACF) AND OF THE CHROMATIN ACCESSIBILITY COMPLEX (CHRAC).
CC       THIS SUBUNIT MAY SERVE AS THE ENERGY-TRANSDUCING COMPONENT OF
CC       CHROMATIN-REMODELING MACHINES.
CC   -!- SUBUNIT: NURF IS COMPOSED OF FOUR SUBUNITS; A 215 kDa PROTEIN,
CC       IMITATION SWITCH (ISWI), NURF-55, AND NURF-38.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY. SNF2L
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L27127; AAA19868.1; -.
DR   EMBL; AE003821; AAF58479.1; -.
DR   FlyBase; FBgn0011604; Iswi.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001005; Myb_DNA_bind.
DR   InterPro; IPR000330; SNF2_N.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00176; SNF2_N; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00395; SANT; 2.
KW   Nuclear protein; Helicase; ATP-binding.
FT   NP_BIND     153    160       ATP (POTENTIAL).
FT   SITE        256    259       DEAH BOX.
FT   DOMAIN      978    981       POLY-LYS.
FT   DOMAIN     1023   1027       POLY-LYS.
SQ   SEQUENCE   1027 AA;  118873 MW;  008FC81AE15E071F CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1027;
  Best Local Similarity   66.7%;  Pred. No. 8.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
             |||  |
Db     658 EAGTSS 663


RESULT  28
PMA1_DICDI
ID   PMA1_DICDI     STANDARD;      PRT;  1058 AA.
AC   P54679;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable plasma membrane ATPase (EC 3.6.3.6) (Proton pump) (PAT2).
GN   PATB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=98083743; PubMed=9421912;
RA   Coukell M.B., Moniakis J., Cameron A.M.;
RT   "The patB gene of Dictyostelium discoideum encodes a P-type H(+)-
RT   ATPase isoform essential for growth and development under acidic
RT   conditions.";
RL   Microbiology 143:3877-3888(1997).
CC   -!- FUNCTION: THE PLASMA MEMBRANE ATPASE IS A HYDROGEN ION PUMP. THE
CC       PROTON GRADIENT IT GENERATES DRIVES THE ACTIVE TRANSPORT OF
CC       NUTRIENTS BY H+- SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION
CC       AND/OR INTERNAL ALKINIZATION MAY MEDIATE GROWTH RESPONSES.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(IN) = ADP + PHOSPHATE +
CC       H(+)(OUT).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC       (E1-E2 ATPASES). SUBFAMILY IIIA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98286; CAA66931.1; -.
DR   DictyDb; DD00061; patB.
DR   InterPro; IPR004014; Cation_ATPase.
DR   InterPro; IPR001757; E1-E2_ATPase.
DR   InterPro; IPR000695; HATPase.
DR   InterPro; IPR001454; Hydrolase.
DR   Pfam; PF00690; Cation_ATPase_N; 1.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PRINTS; PR00120; HATPASE.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Hydrolase; Hydrogen ion transport; Transmembrane; Phosphorylation;
KW   Magnesium; ATP-binding.
FT   DOMAIN        1    212       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    213    232       POTENTIAL.
FT   DOMAIN      233    237       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    238    258       POTENTIAL.
```

```
FT   DOMAIN      259    387       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    388    407       POTENTIAL.
FT   DOMAIN      408    417       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    426    447       POTENTIAL.
FT   DOMAIN      448    783       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    784    805       POTENTIAL.
FT   DOMAIN      806    810       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    811    833       POTENTIAL.
FT   DOMAIN      834    849       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    850    870       POTENTIAL.
FT   DOMAIN      871    889       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    890    910       POTENTIAL.
FT   DOMAIN      911    922       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    923    943       POTENTIAL.
FT   DOMAIN      944    967       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    968    988       POTENTIAL.
FT   DOMAIN      989   1058       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES     480    480       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL       728    728       MAGNESIUM (BY SIMILARITY).
FT   METAL       732    732       MAGNESIUM (BY SIMILARITY).
FT   DOMAIN       44     55       POLY-GLN.
FT   DOMAIN      113    116       POLY-SER.
FT   DOMAIN      246    249       POLY-LEU.
SQ   SEQUENCE   1058 AA;  117373 MW;  CB0E5AB9EDEB9AF2 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1058;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     110 EAGSSS 115


RESULT  29
ACAA_ARATH
ID   ACAA_ARATH     STANDARD;      PRT;  1069 AA.
AC   Q9SZR1; Q9M0D3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Potential calcium-transporting ATPase 10, plasma membrane-type
DE   (EC 3.6.3.8) (Ca2+-ATPase, isoform 10).
GN   ACA10 OR AT4G29900 OR F27B13.140.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC       OF ATP COUPLED WITH THE TRANSLOCATION OF CALCIUM FROM THE CYTOSOL
CC       INTO THE ENDOPLASMIC RETICULUM (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(CIS) = ADP + PHOSPHATE +
CC       CA(2+)(TRANS).
CC   -!- ENZYME REGULATION: ACTIVATED BY CALMODULIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- DOMAIN: THE N-TERMINUS CONTAINS AN AUTOINHIBITORY CALMODULIN-
CC       BINDING DOMAIN, WHICH BINDS CALMODULIN IN A CALCIUM-DEPENDENT
CC       FASHION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC       (E1-E2 ATPASES). SUBFAMILY IIB.
CC   -!- CAUTION: THE SEQUENCE CAB43665 DIFFERS FROM THAT SHOWN DUE TO
CC       WRONG EXON BOUNDARIES PREDICTED FROM THE GENOMIC SEQUENCE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL161575; CAB79748.1; -.
DR   EMBL; AL050352; CAB43665.1; ALT_SEQ.
DR   HSSP; P04191; 1EUL.
DR   InterPro; IPR004014; Cation_ATPase.
DR   InterPro; IPR001757; E1-E2_ATPase.
DR   InterPro; IPR001454; Hydrolase.
DR   Pfam; PF00689; Cation_ATPase_C; 1.
DR   Pfam; PF00690; Cation_ATPase_N; 1.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
KW   ATP-binding; Metal-binding; Magnesium; Calmodulin-binding;
KW   Multigene family.
FT   DOMAIN        1    180       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    181    201       POTENTIAL.
FT   DOMAIN      202    219       LUMENAL (POTENTIAL).
FT   TRANSMEM    220    240       POTENTIAL.
FT   DOMAIN      241    369       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    370    389       POTENTIAL.
FT   DOMAIN      390    426       LUMENAL (POTENTIAL).
FT   TRANSMEM    427    444       POTENTIAL.
```

```
FT   DOMAIN      445    844       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    845    863       POTENTIAL.
FT   DOMAIN      864    874       LUMENAL (POTENTIAL).
FT   TRANSMEM    875    895       POTENTIAL.
FT   DOMAIN      896    915       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    916    938       POTENTIAL.
FT   DOMAIN      939    950       LUMENAL (POTENTIAL).
FT   TRANSMEM    952    973       POTENTIAL.
FT   DOMAIN      974    991       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    992   1013       POTENTIAL.
FT   DOMAIN     1014   1023       LUMENAL (POTENTIAL).
FT   TRANSMEM   1024   1045       POTENTIAL.
FT   DOMAIN     1046   1069       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42     53       CALMODULIN-BINDING (BY SIMILARITY).
FT   MOD_RES     482    482       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL       789    789       MAGNESIUM (BY SIMILARITY).
FT   METAL       793    793       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   1069 AA;  116858 MW;  48CE4A4B218E2656 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1069;
  Best Local Similarity   66.7%;  Pred. No. 8.9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     17 EAGTSS 22


RESULT  30
CNA_STAAU
ID   CNA_STAAU      STANDARD;      PRT;  1183 AA.
AC   Q53654;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Collagen adhesin precursor.
GN   CNA.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FDA 574;
RX   MEDLINE=92165839; PubMed=1311320;
RA   Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,
RA   Lindberg M., Hoeoek M.;
RT   "Molecular characterization and expression of a gene encoding a
RT   Staphylococcus aureus collagen adhesin.";
RL   J. Biol. Chem. 267:4766-4772(1992).
RN   [2]
RP   ERRATUM.
RA   Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,
RA   Lindberg M., Hoeoek M.;
RL   J. Biol. Chem. 269:11672-11672(1994).
RN   [3]
RP   COLLAGEN-BINDING DOMAIN.
RC   STRAIN=FDA 574;
RX   MEDLINE=94032261; PubMed=8218209;
RA   Patti J.M., Boles J.O., Hoeoek M.;
RT   "Identification and biochemical characterization of the ligand
RT   binding domain of the collagen adhesin from Staphylococcus aureus.";
RL   Biochemistry 32:11428-11435(1993).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 169-318.
RX   MEDLINE=97475225; PubMed=9334749;
RA   Symersky J., Patti J.M., Carson M., House-Pompeo K., Teale M.,
RA   Moore D., Jin L., Schneider A., DeLucas L.J., Hoeoek M.,
RA   Narayana S.V.L.;
RT   "Structure of the collagen-binding domain from a Staphylococcus
RT   aureus adhesin.";
RL   Nat. Struct. Biol. 4:833-838(1997).
CC   -!- FUNCTION: MEDIATES ATTACHMENT OF STAPHYOCOCCAL CELLS TO
CC       COLLAGEN-CONTAINING SUBSTRATA.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC       IN THE REGION OF THE MEMBRANE ANCHOR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81736; AAA20874.1; -.
DR   PDB; 1AMX; 24-JUN-98.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; FALSE_NEG.
KW   Signal; Repeat; Transmembrane; Cell wall; 3D-structure.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30   1183       COLLAGEN ADHESIN.
FT   DOMAIN       30   1157       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1158   1177       MEMBRANE ANCHOR (POTENTIAL).
FT   DOMAIN     1178   1183       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      151    318       COLLAGEN-BINDING.
FT   DOMAIN      533   1093       3 X 187 AA APPROXIMATE TANDEM REPEATS.
FT   DOMAIN     1093   1157       LYS/PRO-RICH (CELL WALL-SPANNING).
FT   DOMAIN     1151   1156       CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT                                PROTEINS.
FT   REPEAT      533    719       B1.
FT   REPEAT      720    906       B2.
FT   REPEAT      907   1093       B3.
SQ   SEQUENCE   1183 AA;  133066 MW;  B6A1CC072E575D76 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1183;
  Best Local Similarity   66.7%;  Pred. No. 9.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    166 EAGTSS 171


RESULT  31
B3A2_MOUSE
ID   B3A2_MOUSE     STANDARD;      PRT;  1237 AA.
AC   P13808; Q9ES13; Q9ES12; Q9ES11; Q9ES10; Q9ES09;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3RP).
GN   SLC4A2 OR AE2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=89034212; PubMed=3182834;
RA   Alper S.L., Kopito R.R., Libresco S.M., Lodish H.F.;
RT   "Cloning and characterization of a murine band 3-related cDNA from
RT   kidney and from a lymphoid cell line.";
RL   J. Biol. Chem. 263:17092-17099(1988).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX   MEDLINE=20462926; PubMed=11006093;
RA   Lecanda J., Urtasun R., Medina J.F.;
RT   "Molecular cloning and genomic organization of the mouse AE2 anion
RT   exchanger gene.";
RL   Biochem. Biophys. Res. Commun. 276:117-124(2000).
CC   -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
CC       DISTRIBUTION.
```

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 5 isoforms; a (shown here), b1, b2, c1 and
CC       c2; are produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Isoform a is widely expressed at similar
CC       levels in all tissues examined. Isoforms b1 and b2 are
CC       predominantly expressed in stomach although they are also detected
CC       at lower levels in other tissues. Isoform c1 is stomach-specific.
CC       Isoform c2 is expressed at slightly higher levels in lung and
CC       stomach than in other tissues.
CC   -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04036; AAA65505.1; -.
DR   EMBL; AF255774; AAG23154.1; -.
DR   EMBL; AF255774; AAG23155.1; -.
DR   EMBL; AF255774; AAG23156.1; -.
DR   EMBL; AF255774; AAG23158.1; -.
DR   EMBL; AF255774; AAG23157.1; -.
DR   PIR; A31789; A31789.
DR   HSSP; P02730; 1BTR.
DR   MGD; MGI:109351; Slc4a2.
DR   InterPro; IPR001717; Anion_exchanger.
DR   InterPro; IPR003020; HCO3_cotransp.
DR   Pfam; PF00955; HCO3_cotransp; 1.
DR   PRINTS; PR01231; HCO3TRNSPORT.
DR   PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR   PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW   Transmembrane; Glycoprotein; Anion exchange; Lipoprotein; Palmitate;
KW   Alternative splicing.
FT   DOMAIN        1    703       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      704   1237       MEMBRANE (ANION EXCHANGE).
FT   TRANSMEM    704    727       POTENTIAL.
FT   TRANSMEM    733    770       POTENTIAL.
FT   TRANSMEM    790    812       POTENTIAL.
FT   TRANSMEM    822    843       POTENTIAL.
FT   DOMAIN      844    896       EXOPLASMIC LOOP (POTENTIAL).
FT   TRANSMEM    897    914       POTENTIAL.
FT   DOMAIN      915    929       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    930    950       POTENTIAL.
FT   TRANSMEM    984   1006       POTENTIAL.
FT   TRANSMEM   1032   1053       POTENTIAL.
FT   TRANSMEM   1087   1132       POTENTIAL.
FT   TRANSMEM   1159   1195       POTENTIAL.
FT   DOMAIN        5    316       PRO-RICH.
FT   DOMAIN       73     87       HIS-RICH.
FT   DOMAIN      861    865       POLY-SER.
FT   CARBOHYD    855    855       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    866    866       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    878    878       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID      1169   1169       PALMITATE (BY SIMILARITY).
FT   VARSPLIC      1     17       MSSAPRRPASGADSLHT -> MDFLLRPQ (IN ISOFORM
FT                                B2).
FT   VARSPLIC      1     17       MSSAPRRPASGADSLHT -> MTQ (IN ISOFORM B1).
FT   VARSPLIC      1    166       MISSING (IN ISOFORM C2).
FT   VARSPLIC      1    198       MISSING (IN ISOFORM C1).
FT   VARSPLIC    167    193       ERTSPSPPTQTPHQEAAPRASKGAQTG -> MPAFQEWKSG
FT                                GLREEAVFGAHGCSVCR (IN ISOFORM C2).
FT   CONFLICT    205    205       A -> G (IN REF. 2).
SQ   SEQUENCE   1237 AA;  136813 MW;  1A0782C0071782EE CRC64;


  Query Match           90.5%;  Score 19;  DB 1;  Length 1237;
  Best Local Similarity  66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    858 EAGSSS 863


RESULT  32
MYO6_HUMAN
ID   MYO6_HUMAN     STANDARD;      PRT;  1262 AA.
AC   Q9UM54;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Myosin VI.
GN   MYO6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97402203; PubMed=9259267;
RA   Avraham K.B., Hasson T., Sobe T., Balsara B., Testa J.R.,
RA   Skvorak A.B., Morton C.C., Copeland N.G., Jenkins N.A.;
RT   "Characterization of unconventional MYO6, the human homologue of the
RT   gene responsible for deafness in Snell's waltzer mice.";
RL   Hum. Mol. Genet. 6:1225-1231(1997).
RN   [2]
RP   REVISIONS.
RA   Avraham K.B.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   FUNCTION.
RX   MEDLINE=99447046; PubMed=10519557;
RA   Wells A.L., Lin A.W., Chen L.-Q., Safer D., Cain S.M., Hasson T.,
RA   Carragher B.O., Milligan R.A., Sweeney H.L.;
RT   "Myosin VI is an actin-based motor that moves backwards.";
RL   Nature 401:505-508(1999).
RN   [4]
RP   VARIANT DFNA22 TYR-442.
RX   MEDLINE=21375673; PubMed=11468689;
RA   Melchionda S., Ahituv N., Bisceglia L., Sobe T., Glaser F.,
RA   Rabionet R., Arbones M.L., Notarangelo A., Di Iorio E., Carella M.,
RA   Zelante L., Estivill X., Avraham K.B., Gasparini P.;
RT   "MYO6, the human homologue of the gene responsible for deafness in
RT   Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic
RT   hearing loss.";
RL   Am. J. Hum. Genet. 69:635-640(2001).
CC   -!- FUNCTION: RECESSIVE ACTIN-BASED MOTOR. REQUIRED FOR STRUCTURAL
CC       INTEGRITY OF INNER EAR HAIR CELLS (BY SIMILARITY).
CC   -!- DISEASE: Defects in MYO6 are the cause of an autosomal dominant
CC       form of nonsyndromic sensorineural deafness (DFNA22). The deafness
CC       is progressive and postlingual, with onset during childhood (8 to
CC       10 years of age at onset of symptoms; 6 to 8 years of age at onset
CC       of first audiometric abnormalities). By the age of approximately
CC       50 years, affected individuals invariably have profound
CC       sensorineural deafness.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U90236; AAC51654.2; -.
DR   HSSP; P08799; 1LVK.
DR   MIM; 600970; -.
DR   MIM; 606346; -.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
```

```
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; FALSE_NEG.
KW   Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Coiled coil; Disease mutation; Deafness.
FT   DOMAIN        1    759       MYOSIN HEAD-LIKE.
FT   DOMAIN      814    834       IQ.
FT   DOMAIN      848   1030       COILED COIL (POTENTIAL).
FT   NP_BIND     151    158       ATP (POTENTIAL).
FT   DOMAIN      665    672       ACTIN-BINDING (POTENTIAL).
FT   VARIANT     442    442       C -> Y (IN DFNA22).
FT                                /FTId=VAR_012110.
SQ   SEQUENCE   1262 AA;  146047 MW;  CF1FA35796FC1C60 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1262;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    354 EAGSTS 359


RESULT  33
DYNA_DROME
ID   DYNA_DROME     STANDARD;      PRT;  1265 AA.
AC   P13496; Q9VUA1;
DT   01-JAN-1990 (Rel. 13, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   150 kDa dynein-associated polypeptide (DP-150) (DAP-150) (Glued
DE   protein).
GN   GL OR CG9206.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R, AND CANTON-S;
RX   MEDLINE=87317680; PubMed=2819881;
RA   Swaroop A., Swaroop M., Garen A.;
RT   "Sequence analysis of the complete cDNA and encoded polypeptide for
RT   the Glued gene of Drosophila melanogaster.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:6501-6505(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: REQUIRED FOR THE CYTOPLASMIC DYNEIN-DRIVEN RETROGRADE
CC       MOVEMENT OF VESICLES AND ORGANELLES ALONG MICROTUBULES. DYNEIN-
CC       DYNACTIN INTERACTION IS A KEY COMPONENT OF THE MECHANISM OF AXONAL
CC       TRANSPORT OF VESICLES AND ORGANELLES.
CC   -!- SUBUNIT: LARGE MACROMOLECULAR COMPLEX OF AT LEAST 10 COMPONENTS.
CC       P150(GLUED) BINDS DIRECTLY TO MICROTUBULES AND TO CYTOPLASMIC
CC       DYNEIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS THE DYNACTIN 150 KDA SUBUNIT FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
CC   -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC       FRAMESHIFTS AT POSITIONS 32; 174 TO 220; 648 TO 672 AND 1208.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02932; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE003536; AAF49788.1; -.
DR   PIR; A28313; A28313.
DR   FlyBase; FBgn0001108; Gl.
DR   InterPro; IPR000938; CAP-Gly.
DR   Pfam; PF01302; CAP_GLY; 1.
DR   PROSITE; PS00845; CAP_GLY_1; 1.
DR   PROSITE; PS50245; CAP_GLY_2; 1.
KW   Motor protein; Microtubules; Dynein; Coiled coil; Cytoskeleton.
FT   DOMAIN       27     69       CAP-GLY.
FT   DOMAIN      105    138       SER-RICH.
FT   DOMAIN      213    570       COILED COIL (POTENTIAL).
FT   DOMAIN      812    836       COILED COIL (POTENTIAL).
FT   DOMAIN      967   1084       COILED COIL (POTENTIAL).
FT   DOMAIN     1128   1160       COILED COIL (POTENTIAL).
FT   CONFLICT    708    708       D -> A (IN REF. 1).
FT   CONFLICT    875    875       L -> V (IN REF. 1).
FT   CONFLICT    888    888       A -> R (IN REF. 1).
FT   CONFLICT   1043   1043       S -> C (IN REF. 1).
SQ   SEQUENCE   1265 AA;  141217 MW;  2038A200282B2755 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1265;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    802 EAGATS 807


RESULT  34
MYO6_MOUSE
```

```
ID   MYO6_MOUSE     STANDARD;      PRT;  1265 AA.
AC   Q64331;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Myosin VI.
GN   MYO6 OR SV.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96083582; PubMed=7493015;
RA   Avraham K.B., Hasson T., Steel K.P., Kingsley D.M., Russell L.B.,
RA   Mooseker M.S., Copeland N.G., Jenkins N.A.;
RT   "The mouse Snell's waltzer deafness gene encodes an unconventional
RT   myosin required for structural integrity of inner ear hair cells.";
RL   Nat. Genet. 11:369-375(1995).
CC   -!- FUNCTION: RECESSIVE ACTIN-BASED MOTOR. REQUIRED FOR STRUCTURAL
CC       INTEGRITY OF INNER EAR HAIR CELLS (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. WITHIN THE COCHLEA,
CC       EXPRESSED SPECIFICALLY WITHIN THE SENSORY HAIR CELLS.
CC   -!- DISEASE: DEFECTS IN MYO6 ARE THE CAUSE OF SNELL'S WALTZER, A
CC       CONDITION CHARACTERIZED BY CIRCLING, HEAD-TOSSING, DEAFNESS AND
CC       HYPERACTIVITY.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49739; AAB00194.1; -.
DR   HSSP; P08799; 1LVK.
DR   MGD; MGI:104785; Myo6.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 4.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; FALSE_NEG.
KW   Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Coiled coil; Disease mutation; Deafness.
FT   DOMAIN        1    762       MYOSIN HEAD-LIKE.
FT   DOMAIN      817    837       IQ.
FT   DOMAIN      849   1014       COILED COIL (POTENTIAL).
FT   NP_BIND     151    158       ATP (POTENTIAL).
FT   DOMAIN      668    675       ACTIN-BINDING (POTENTIAL).
FT   VARIANT     766   1265       MISSING (IN SNELL'S WALTZER).
SQ   SEQUENCE   1265 AA;  146408 MW;  4F51ABC72463148C CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1265;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    355 EAGSTS 360


RESULT  35
Y232_HUMAN
ID   Y232_HUMAN     STANDARD;      PRT;  1278 AA.
AC   Q92628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein KIAA0232 (Fragment).
GN   KIAA0232.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA   Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 271-1278 FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=97191544; PubMed=9039502;
RA   Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA   Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. VI.
RT   The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT   analysis of cDNA clones from cell line KG-1 and brain.";
RL   DNA Res. 3:321-329(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D86985; BAA13221.2; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   1278 AA;  141663 MW;  2FCFC8837AF8134D CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1278;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    177 EAGSSS 182


RESULT  36
CUT1_MOUSE
ID   CUT1_MOUSE     STANDARD;      PRT;  1395 AA.
AC   P53564; O08994; P70301;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   CCAAT displacement protein (CDP) (Cut-like 1) (Homeobox protein Cux)
DE   (Fragment).
GN   CUTL1 OR CUX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6N;
RA   Wang Z., Goldstein A., Neufeld E.J., Scheuermann R.H., Tucker P.W.;
RT   "Repression of immunoglobulin heavy chain intronic enhancer
RT   through nuclear matrix attachment sites: Cux/CDP homeoprotein is a
RT   component of NF-muNR repressor.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).
```

```
RC   STRAIN=A/J, AND BALB/C; TISSUE=Brain;
RX   MEDLINE=94244481; PubMed=7910552;
RA   Valarche I., Tissier-Seta J.P., Hirsch M.R., Martinez S., Goridis C.,
RA   Brunet J.F.;
RT   "The mouse homeodomain protein Phox2 regulates Ncam promoter activity
RT   in concert with Cux/CDP and is a putative determinant of
RT   neurotransmitter phenotype.";
RL   Development 119:881-896(1993).
RN   [3]
RP   SEQUENCE OF 642-1395 FROM N.A.
RX   MEDLINE=96437626; PubMed=8840273;
RA   den Heuvel G.B., Bodmer R., McConnell K.R., Nagami G.T., Igarashi P.;
RT   "Expression of a cut-related homeobox gene in developing and
RT   polycystic mouse kidney.";
RL   Kidney Int. 50:453-461(1996).
RN   [4]
RP   SEQUENCE OF 936-1395 FROM N.A.
RC   TISSUE=Testis;
RA   Quaggin S.E., Igarashi P.;
RT   "A unique variant of a homeobox gene related to Drosophila cut is
RT   expressed in mouse testis.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
CC       REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
CC       PREVENTING BINDING OF POSITIVELY-ACTIVING CCAAT FACTORS TO
CC       PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR; BINDS
CC       TO THE MARS (5' AND 3') OF THE IMMUNOGLOBULIN HEAVY CHAIN
CC       ENHANCER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEOBOX PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF004225; AAD12485.1; -.
DR   EMBL; X75013; CAA52922.1; -.
DR   EMBL; U46683; AAC52775.1; -.
DR   EMBL; U46684; AAB41146.1; -.
DR   HSSP; P10037; 1AU7.
DR   MGD; MGI:88568; Cutl1.
DR   InterPro; IPR003350; CUT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF02376; CUT; 3.
DR   Pfam; PF00046; homeobox; 2.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Transcription regulation; Homeobox; DNA-binding;
KW   Developmental protein; Nuclear protein; Repeat; Repressor;
KW   Coiled coil; Alternative splicing.
FT   NON_TER       1      1
FT   DOMAIN        1    243       COILED COIL (POTENTIAL).
FT   DNA_BIND    420    507       CUT 1.
FT   DOMAIN      547    603       COILED COIL (POTENTIAL).
FT   DNA_BIND    809    896       CUT 2.
FT   DNA_BIND    992   1079       CUT 3.
FT   DNA_BIND   1119   1178       HOMEOBOX.
FT   VARSPLIC    287    388       MISSING (IN ISOFORM 2).
FT   CONFLICT   1360   1360       G -> A (IN REF. 2).
FT   CONFLICT   1365   1365       P -> L (IN REF. 1).
SQ   SEQUENCE   1395 AA;  151802 MW;  D062CC227D7A163E CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1395;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+03;
    Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
             |||  |
Db     377 EAGSTS 382


RESULT  37
MUKB_ECOLI
ID   MUKB_ECOLI     STANDARD;      PRT;  1486 AA.
AC   P22523; P77164; Q47398;
DT   01-AUG-1991 (Rel. 19, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cell division protein mukB.
GN   MUKB OR B0924.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=91114703; PubMed=1989883;
RA   Niki H., Jaffe A., Imamura R., Ogura T., Hiraga S.;
RT   "The new gene mukB codes for a 177 kd protein with coiled-coil
RT   domains involved in chromosome partitioning of E. coli.";
RL   EMBO J. 10:183-193(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND MUTANTS MUKB33 AND MUKB106.
RX   MEDLINE=95080615; PubMed=7988894;
RA   Yamanaka K., Mitani T., Feng J., Ogura T., Niki H., Hiraga S.;
RT   "Two mutant alleles of mukB, a gene essential for chromosome
RT   partition in Escherichia coli.";
RL   FEMS Microbiol. Lett. 123:27-31(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [5]
RP   SEQUENCE OF 1-44 FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=94232180; PubMed=7513784;
RA   Feng J., Yamanaka K., Niki H., Ogura T., Hiraga S.;
RT   "New killing system controlled by two genes located immediately
RT   upstream of the mukB gene in Escherichia coli.";
RL   Mol. Gen. Genet. 243:136-147(1994).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-227.
RX   MEDLINE=20015369; PubMed=10545328;
RA   van den Ent F., Lockhart A., Kendrick-Jones J., Loewe J.;
RT   "Crystal structure of the N-terminal domain of MukB: a protein
RT   involved in chromosome partitioning.";
RL   Structure 7:1181-1187(1999).
```

```
CC   -!- FUNCTION: ESSENTIAL FOR CHROMOSOME PARTITIONING. IMPLICATED IN
CC       ATP-DEPENDENT CHROMOSOME PARTITIONING DURING CELL DIVISION.
CC   -!- SIMILARITY: CONTAINS A COILED COIL MYOSIN-LIKE STRUCTURE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X57550; CAA40776.1; -.
DR   EMBL; D31701; BAA06510.1; -.
DR   EMBL; AE000194; AAC74010.1; -.
DR   EMBL; D90730; BAA35670.1; -.
DR   EMBL; D26440; BAA05459.1; -.
DR   PIR; JH0228; JH0228.
DR   PDB; 1QHL; 10-NOV-99.
DR   EcoGene; EG10618; mukB.
KW   ATP-binding; Coiled coil; 3D-structure; Complete proteome.
FT   DOMAIN      331    665       COILED COIL (POTENTIAL).
FT   DOMAIN      784   1116       COILED COIL (POTENTIAL).
FT   DOMAIN     1209   1265       COILED COIL (POTENTIAL).
FT   NP_BIND      34     41       ATP (BY SIMILARITY).
FT   VARIANT      33     33       S -> F (IN MUKB106).
FT   VARIANT    1201   1201       D -> N (IN MUKB33).
FT   CONFLICT    266    266       A -> R (IN REF. 1 AND 2).
FT   CONFLICT    318    319       EH -> DD (IN REF. 1).
FT   CONFLICT   1134   1134       H -> D (IN REF. 1).
FT   CONFLICT   1174   1175       SE -> VQ (IN REF. 1).
FT   CONFLICT   1276   1277       MISSING (IN REF. 1 AND 2).
FT   CONFLICT   1357   1380       WLRAESGALSTGEAIGTGMSILVM -> CCAQSLVHCRPVR
FT                                RWYRYVDSGV (IN REF. 1).
FT   CONFLICT   1389   1486       SRRLRGKDISPCRLLFLDEAARLDARSIATLFELCERLQMQ
FT                                LIIAAPENISPEKGTTYKLVRKVFQNTEHVHVVGLRGFAPQ
FT                                LPETLPGTDEAPSQAS -> SAACAVKISLLAACCSSMKQR
FT                                DWMLVLSPRCLNCVSVCKCNSSSQRRKISARRKAPPINWCV
FT                                KSSRIPNTFMSSACEDLRRNSLKRFQELTKRLLRRVKIKQQ
FT                                CRLFFFRKLRFCTKKVAHYGALFFKLLYIRLCKNVRRLYTE
FT                                DKPDE (IN REF. 1).
SQ   SEQUENCE   1486 AA;  170229 MW;  38C7874BEB78D6D6 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1486;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
            |||  |
Db     65 EAGATS 70


RESULT  38
CUT1_HUMAN
ID   CUT1_HUMAN     STANDARD;      PRT;  1505 AA.
AC   P39880; Q9UEV5;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   CCAAT displacement protein (CDP) (Cut-like 1).
GN   CUTL1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical vein;
RX   MEDLINE=93250973; PubMed=1301999;
RA   Neufeld E.J., Skalnik D.G., Lievens P.M.-J., Orkin S.H.;
RT   "Human CCAAT displacement protein is homologous to the Drosophila
RT   homeoprotein, cut.";
RL   Nat. Genet. 1:50-55(1992).
RN   [2]
RP   SEQUENCE OF 48-224 FROM N.A.
RX   MEDLINE=99018118; PubMed=9799793;
RA   Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA   Tsui L.-C., Rosenthal A.;
RT   "Large-scale sequencing of two regions in human chromosome 7q22:
RT   analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT   reveals 17 genes.";
RL   Genome Res. 8:1060-1073(1998).
CC   -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
CC       REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
CC       PREVENTING BINDING OF POSITIVELY-ACTIVING CCAAT FACTORS TO
CC       PROMOTERS.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEOBOX MAY PARTICIPATE IN
CC       REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC       HETERODIMERIZATION.
CC   -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEOBOX PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74099; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF047825; AAC78778.1; -.
DR   HSSP; P10037; 1AU7.
DR   TRANSFAC; T00100; -.
DR   MIM; 116896; -.
DR   InterPro; IPR003350; CUT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF02376; CUT; 3.
DR   Pfam; PF00046; homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Transcription regulation; Homeobox; DNA-binding;
KW   Developmental protein; Nuclear protein; Repeat; Repressor;
KW   Coiled coil; Alternative splicing.
FT   DOMAIN        7    363       COILED COIL (POTENTIAL).
FT   DNA_BIND    542    629       CUT 1.
FT   DOMAIN      669    725       COILED COIL (POTENTIAL).
FT   DNA_BIND    934   1021       CUT 2.
FT   DNA_BIND   1117   1204       CUT 3.
FT   DNA_BIND   1244   1303       HOMEOBOX.
FT   VARSPLIC    632    653       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   1505 AA;  164353 MW;  860E14D508D4DE11 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1505;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
            |||  |
Db    499 EAGSTS 504


RESULT  39
UGGG_DROME
ID   UGGG_DROME     STANDARD;      PRT;  1548 AA.
AC   Q09332;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   UDP-glucose:glycoprotein glucosyltransferase precursor (EC 2.4.1.-)
```

```
DE   (UDP-Glc:glycoprotein glucosyltransferase) (dUGT).
GN   UGT OR UGGG.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-37.
RC   TISSUE=Embryo;
RX   MEDLINE=95246722; PubMed=7729408;
RA   Parker C.G., Fessler L.I., Nelson R.E., Fessler J.H.;
RT   "Drosophila UDP-glucose:glycoprotein glucosyltransferase: sequence
RT   and characterization of an enzyme that distinguishes between
RT   denatured and native proteins.";
RL   EMBO J. 14:1294-1303(1995).
CC   -!- FUNCTION: UNFOLDED, DENATURED GLYCOPROTEINS ARE SUBSTANTIALLY
CC       BETTER SUBSTRATES FOR GLUCOSYLATION BY THIS ENZYME THAN ARE THE
CC       CORRESPONDING NATIVE PROTEINS. THIS PROTEIN AND TRANSIENT
CC       GLUCOSYLATION MAY BE INVOLVED IN MONITORING AND/OR ASSISTING THE
CC       FOLDING AND ASSEMBLY OF NEWLY MADE GLYCOPROTEINS, IN ORDER TO
CC       IDENTIFY GLYCOPROTEINS THAT NEED ASSISTANCE IN FOLDING FROM
CC       CHAPERONES.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MANGANESE IONS FOR ACTIVITY.
CC   -!- PATHWAY: GLYCOSYLATION.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum.
CC   -!- DEVELOPMENTAL STAGE: IS PRESENT AT LOW BUT DETECTABLE LEVELS IN
CC       THE EARLIEST EMBRYOS, INCREASING AT 6-8 HRS WITH A MAXIMUM AT 10-
CC       12 HRS. LEVELS DECREASE THEREAFTER AND ARE NOT DETECTED IN 18-20
CC       HRS EMBRYOS AND FIRST INSTAR LARVAE BUT IS DETECTED AGAIN AT
CC       SECOND INSTAR TO PUPATION.
CC   -!- SIMILARITY: SOME, TO YEAST KRE5, S.TYPHIMURIUM RFAJ AND E.COLI
CC       RFAI PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U20554; AAA85850.1; -.
DR   FlyBase; FBgn0014075; Ugt.
DR   InterPro; IPR002495; Glycosyl_transf_8.
DR   Pfam; PF01501; Glyco_transf_8; 1.
KW   Signal; Transferase; Glycosyltransferase; Endoplasmic reticulum;
KW   Glycoprotein.
FT   SIGNAL        1     22
FT   CHAIN        23   1548       UDP-GLUCOSE:GLYCOPROTEIN
FT                                GLUCOSYLTRANSFERASE.
FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    864    864       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SITE       1545   1548       PREVENT SECRETION FROM ER (POTENTIAL).
SQ   SEQUENCE   1548 AA;  174465 MW;  95D6849961622DB6 CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 1548;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     255 EAGSTS 260


RESULT  40
HMP2_YEREN
ID   HMP2_YEREN     STANDARD;      PRT;  2035 AA.
AC   P48633;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   High-molecular-weight protein 2 (HMWP2).
GN   IRP2.
OS   Yersinia enterocolitica.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Yersinia.
OX   NCBI_TaxID=630;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8081 / SEROTYPE O:8;
RX   MEDLINE=93374844; PubMed=8366034;
RA   Guilvout I., Mercereau-Puijalon O., Bonnefoy S., Pugsley A.P.,
RA   Carniel E.;
RT   "High-molecular-weight protein 2 of Yersinia enterocolitica is
RT   homologous to AngR of Vibrio anguillarum and belongs to a family of
RT   proteins involved in nonribosomal peptide synthesis.";
RL   J. Bacteriol. 175:5488-5504(1993).
CC   -!- FUNCTION: UNKNOWN. MAY BE INVOLVED IN THE NONRIBOSOMAL SYNTHESIS
CC       OF SMALL PEPTIDES.
CC   -!- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC       (POTENTIAL).
CC   -!- DOMAIN: CONSISTS OF A CENTRAL REGION WITH SIMILARITY TO THE REPEAT
CC       DOMAINS OF ACVS AND GRC2, FLANKED BY TWO REPEAT DOMAINS, EACH OF
CC       WHICH CONTAINS 5 DIRECT REPEATS.
CC   -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L18881; AAA27636.1; -.
DR   EMBL; Z35454; CAA84606.1; -.
DR   PIR; A48654; A48654.
DR   HSSP; P14687; 1AMU.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; DUF4.
DR   InterPro; IPR003880; Phosphopant_attach.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00668; Condensation; 2.
DR   Pfam; PF00550; pp-binding; 3.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 3.
KW   Ligase; Multifunctional enzyme; Phosphopantetheine; Repeat.
FT   DOMAIN        3    547       I.
FT   REPEAT      114    146       I-DR1.
FT   REPEAT      310    321       I-DR2.
FT   REPEAT      378    390       I-DR3.
FT   REPEAT      454    462       I-DR4.
FT   REPEAT      477    491       I-DR5.
FT   DOMAIN     1466   1919       II.
FT   REPEAT     1495   1527       II-DR1.
FT   REPEAT     1682   1693       II-DR2.
FT   REPEAT     1750   1762       II-DR3.
FT   REPEAT     1826   1834       II-DR4.
FT   REPEAT     1849   1863       II-DR5.
FT   DOMAIN       20     88       ACYL CARRIER (ACP) 1.
FT   DOMAIN     1409   1475       ACYL CARRIER (ACP) 2.
FT   DOMAIN     1944   2014       ACYL CARRIER (ACP) 3.
FT   BINDING      52     52       PHOSPHOPANTETHEINE (BY SIMILARITY).
SQ   SEQUENCE   2035 AA;  228826 MW;  1C801377A4375BDC CRC64;


  Query Match             90.5%;  Score 19;  DB 1;  Length 2035;
```

```
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db   1972 EAGATS 1977


RESULT  41
CLPA_PINPS
ID   CLPA_PINPS     STANDARD;      PRT;    30 AA.
AC   P81671;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   ATP-dependent clp protease ATP-binding subunit clpA homolog
DE   (Fragments).
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=71647;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Needle;
RX   MEDLINE=99274088; PubMed=10344291;
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- FUNCTION: MAY INTERACT WITH A CLPP-LIKE PROTEASE INVOLVED IN
CC       DEGRADATION OF DENATURED PROTEINS IN THE CHLOROPLAST (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC       (SPOT N9) IS: 5.9, ITS MW IS: 92 kDa.
CC   -!- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
DR   InterPro; IPR001270; CLP_AB.
DR   PROSITE; PS00870; CLPAB_1; PARTIAL.
DR   PROSITE; PS00871; CLPAB_2; PARTIAL.
KW   Chaperone; ATP-binding; Repeat; Chloroplast.
FT   NON_TER       1      1
FT   NON_CONS     15     16
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  2923 MW;  44B5950B73A96152 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 47;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db      4 EAGDAS 9


RESULT  42
PLAS_CAPBU
ID   PLAS_CAPBU     STANDARD;      PRT;    99 AA.
AC   P00294;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Plastocyanin.
GN   PETE.
OS   Capsella bursa-pastoris (Shepherd's purse).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Capsella.
OX   NCBI_TaxID=3719;
RN   [1]
RP   SEQUENCE.
RA   Scawen M.D., Ramshaw J.A.M., Brown R.H., Boulter D.;
RL   Unpublished results, cited by:
RL   Boulter D., Haslett B.G., Peacock D., Ramshaw J.A.M., Scawen M.D.;
RL   (In) Northcote D.H. (eds.);
RL   Plant biochemistry II, pp.13:1-40, University Park Press,
RL   Baltimore (1977).
CC   -!- FUNCTION: PLASTOCYANIN PARTICIPATES IN ELECTRON TRANSFER BETWEEN
CC       P700 AND THE CYTOCHROME B/F COMPLEX IN PHOTOSYSTEM I.
CC   -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
CC       MEMBRANE SURFACE IN CHLOROPLASTS.
CC   -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
DR   PIR; A00304; CUSU.
DR   HSSP; P00289; 2PCF.
DR   Mendel; 11575; CAPbu;PetE;1.
DR   InterPro; IPR001235; Copper_blue.
DR   InterPro; IPR000923; Copper_blue1.
DR   Pfam; PF00127; copper-bind; 1.
DR   PRINTS; PR00156; COPPERBLUE.
DR   ProDom; PD001235; Copper_blue; 1.
DR   PROSITE; PS00196; COPPER_BLUE; 1.
KW   Chloroplast; Electron transport; Copper; Thylakoid; Membrane.
FT   DOMAIN        1     99       PLASTOCYANIN-LIKE.
FT   METAL        37     37       COPPER (BY SIMILARITY).
FT   METAL        84     84       COPPER (BY SIMILARITY).
FT   METAL        87     87       COPPER (BY SIMILARITY).
FT   METAL        92     92       COPPER (BY SIMILARITY).
SQ   SEQUENCE   99 AA;  10383 MW;  30BA97B58B9580F1 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 99;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     76 EAGTYS 81


RESULT  43
WN14_HUMAN
ID   WN14_HUMAN     STANDARD;      PRT;   123 AA.
AC   O14904;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   WNT-14 protein (Fragment).
GN   WNT14.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98110581; PubMed=9441749;
RA   Bergstein I., Eisenberg L.M., Bhalerao J., Jenkins N.A.,
RA   Copeland N.G., Osborne M.P., Bowcock A.M., Brown A.M.C.;
RT   "Isolation of two novel WNT genes, WNT14 and WNT15, one of which
RT   (WNT15) is closely linked to WNT3 on human chromosome 17q21.";
RL   Genomics 46:450-458(1997).
CC   -!- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC       TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A
CC       SIGNALING MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE
CC       REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL
CC       DIAMETERS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC       extracellular matrix.
CC   -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF028702; AAC39550.1; -.
DR   MIM; 602863; -.
DR   InterPro; IPR000970; Wnt1.
DR   Pfam; PF00110; wnt; 1.
DR   SMART; SM00097; WNT1; 1.
DR   PROSITE; PS00246; WNT1; PARTIAL.
KW   Developmental protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER     123    123
SQ   SEQUENCE   123 AA;  13143 MW;  8F000D2568EEA744 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 123;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     43 EAGAIS 48


RESULT  44
PPIB_BACSU
ID   PPIB_BACSU     STANDARD;      PRT;   143 AA.
AC   P35137;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Peptidyl-prolyl cis-trans isomerase B (EC 5.2.1.8) (PPIase B)
DE   (Rotamase B).
GN   PPIB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / MARBURG;
RX   MEDLINE=95020538; PubMed=7934829;
RA   Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;
RT   "The organization of the Bacillus subtilis 168 chromosome region
RT   between the spoVA and serA genetic loci, based on sequence data.";
RL   Mol. Microbiol. 10:385-395(1993).
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=168 / JH642;
RX   MEDLINE=94293776; PubMed=8022278;
RA   Herrler M., Bang H., Marahiel M.A.;
RT   "Cloning and characterization of ppiB, a Bacillus subtilis gene which
RT   encodes a cyclosporin A-sensitive peptidyl-prolyl cis-trans
RT   isomerase.";
RL   Mol. Microbiol. 11:1073-1083(1994).
RN   [3]
RP   SEQUENCE OF 1-26.
RC   STRAIN=168 / JH642;
RX   MEDLINE=96345629; PubMed=8755892;
RA   Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT   "Cold shock stress-induced proteins in Bacillus subtilis.";
RL   J. Bacteriol. 178:4611-4619(1996).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS.
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- ENZYME REGULATION: INHIBITED BY CYCLOSPORIN A (CSA).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PPIASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L09228; AAA67475.1; -.
DR   EMBL; X73898; CAA52103.1; -.
DR   EMBL; Z99116; CAB14268.1; -.
DR   PIR; S45537; S45537.
DR   HSSP; Q27450; 1A58.
DR   SubtiList; BG10512; ppiB.
DR   InterPro; IPR002130; CSA_PPIase.
DR   Pfam; PF00160; pro_isomerase; 1.
DR   PRINTS; PR00153; CSAPPISMRASE.
DR   PROSITE; PS00170; CSA_PPIASE_1; 1.
DR   PROSITE; PS50072; CSA_PPIASE_2; 1.
KW   Isomerase; Rotamase; Complete proteome.
SQ   SEQUENCE   143 AA;  15256 MW;  9EF17D70EB81EC51 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 143;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     82 EAGALS 87


RESULT  45
R157_BOVIN
ID   R157_BOVIN     STANDARD;      PRT;   147 AA.
AC   Q28183;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Retina-specific 15.7 kDa protein.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=86258168; PubMed=2425311;
RA   Nakagawa Y., Kuo C.H., Ishii K., Shiosaka S., Tohyama M., Miki N.;
RT   "Cloning and characterization of a cDNA specific for bovine retina.";
RL   Neurosci. Res. 3:300-310(1986).
CC   -!- TISSUE SPECIFICITY: RETINA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M34915; AAA30756.1; -.
SQ   SEQUENCE   147 AA;  15658 MW;  1FEDA48788B39645 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 147;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    133 EAGTVS 138


RESULT  46
DUT_MYCTU
ID   DUT_MYCTU      STANDARD;      PRT;   154 AA.
```

```
AC   O07199;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
DE   (dUTPase) (dUTP pyrophosphatase).
GN   DUT OR RV2697C OR MT2771 OR MTCY05A6.18C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS ENZYME IS INVOLVED IN NUCLEOTIDE METABOLISM: IT
CC       PRODUCES DUMP, THE IMMEDIATE PRECURSOR OF THYMIDINE NUCLEOTIDES
CC       AND IT DECREASES THE INTRACELLULAR CONCENTRATION OF DUTP SO THAT
CC       URACIL CANNOT BE INCORPORATED INTO DNA (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: dUTP + H(2)O = dUMP + diphosphate.
CC   -!- PATHWAY: DE NOVO SYNTHESIS OF THYMIDYLATE.
CC   -!- SIMILARITY: BELONGS TO THE DUTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z96072; CAB09487.1; -.
DR   EMBL; AE007106; AAK47086.1; -.
DR   HSSP; P06968; 1EUW.
DR   TIGR; MT2771; -.
DR   TubercuList; Rv2697c; -.
DR   InterPro; IPR001428; dUTPase.
DR   Pfam; PF00692; dUTPase; 1.
DR   ProDom; PD000946; dUTPase; 1.
KW   Hydrolase; Nucleotide metabolism; Complete proteome.
SQ   SEQUENCE   154 AA;  15803 MW;  836D5E6420EF455B CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 154;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db     132 EAGLAS 137
```

```
RESULT  47
MOAE_RHIME
ID   MOAE_RHIME     STANDARD;      PRT;   155 AA.
AC   Q92QX5;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Molybdopterin converting factor subunit 2 (MPT synthase subunit 2)
DE   (Molybdopterin synthase subunit 2) (Molybdenum cofactor biosynthesis
DE   protein E) (Molybdopterin converting factor large subunit).
GN   MOAE OR R01168 OR SMC00599.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC   -!- FUNCTION: Converts molybdopterin precursor Z into molybdopterin.
CC       This requires the incorporation of two sulfur atoms into precursor
CC       Z to generate ae dithiolene group (By similarity).
CC   -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC   -!- SUBUNIT: Heterodimer of 2 moaD subunits and 2 moaE subunits (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE MOAE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL591786; CAC45747.1; -.
DR   InterPro; IPR003448; MoaE.
DR   Pfam; PF02391; MoeA; 1.
KW   Molybdenum cofactor biosynthesis; Complete proteome.
SQ   SEQUENCE   155 AA;  16926 MW;  4DE035E7ADFBB367 CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 155;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      43 EAGALS 48


RESULT  48
CEST_ECO57
ID   CEST_ECO57     STANDARD;      PRT;   156 AA.
AC   P58233;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tir chaperone.
GN   CEST OR Z5111 OR ECS4560 OR L0026.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=83334;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / HA1;
RX   MEDLINE=96079490; PubMed=8566710;
RA   Zhao S., Mitchell S.E., Meng J., Doyle M.P., Kresovich S.;
RT   "Cloning and nucleotide sequence of a gene upstream of the eaeA gene
RT   of enterohemorrhagic Escherichia coli O157:H7.";
RL   FEMS Microbiol. Lett. 133:35-39(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=98339885; PubMed=9673266;
RA   Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA   Kaper J.B., Blattner F.R.;
RT   "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT   Escherichia coli O157:H7.";
RL   Infect. Immun. 66:3810-3817(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: Chaperone for the type III secretion of Tir. Probably
CC       stabilizes the protein, prevents inappropriate protein-
CC       proteininteractions and aids in secretion (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE CEST/SYCH CHAPERONE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32312; AAB00110.1; -.
DR   EMBL; AF071034; AAC31505.1; -.
DR   EMBL; AE005595; AAG58824.1; -.
DR   EMBL; AP002566; BAB37983.1; -.
KW   Chaperone; Virulence.
SQ   SEQUENCE   156 AA;  17681 MW;  999545426E26D2D6 CRC64;
```

```
  Query Match           85.7%;  Score 18;  DB 1;  Length 156;
  Best Local Similarity  66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      87 EAGAQS 92
```

```
RESULT  49
FLIN_PSEAE
```

```
ID   FLIN_PSEAE     STANDARD;      PRT;   157 AA.
AC   Q51466;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar motor switch protein fliN.
GN   FLIN OR PA1444.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAK;
RX   MEDLINE=95347807; PubMed=7622217;
RA   Simpson D.A., Ramphal R., Lory S.;
RT   "Characterization of Pseudomonas aeruginosa fliO, a gene involved in
RT   flagellar biosynthesis and adherence.";
RL   Infect. Immun. 63:2950-2957(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- FUNCTION: FLIN IS ONE OF THREE PROTEINS (FLIG, FLIN, FLIM) THAT
CC       FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE
CC       OF THE BASAL BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ
CC       CHEMOTAXIS PROTEINS, IN ADDITION TO CONTACTING COMPONENTS OF THE
CC       MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FLIN/MOPA/SPAO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L39832; AAA79754.1; -.
DR   EMBL; AE004574; AAG04833.1; -.
DR   InterPro; IPR001172; Flag_FliN.
DR   InterPro; IPR001543; SpoA.
DR   Pfam; PF01052; SpoA; 1.
DR   PRINTS; PR00956; FLGMOTORFLIN.
DR   ProDom; PD001777; SpoA; 1.
KW   Chemotaxis; Flagella; Flagellar rotation; Inner membrane;
KW   Complete proteome.
FT   CONFLICT     48     48       P -> S (IN REF. 1).
SQ   SEQUENCE   157 AA;  16620 MW;  B3D91C0182ACB775 CRC64;
```

```
  Query Match           85.7%;  Score 18;  DB 1;  Length 157;
  Best Local Similarity  66.7%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
           |||  |
Db      24 EAGDAS 29
```

```
RESULT  50
YMH2_CAEEL
```

```
ID   YMH2_CAEEL     STANDARD;      PRT;   159 AA.
AC   P34469;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Hypothetical 17.2 kDa protein F58A4.2 in chromosome III.
GN   F58A4.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z22179; CAA80168.1; -.
DR   PIR; S40974; S40974.
DR   WormPep; F58A4.2; CE01017.
KW   Hypothetical protein.
SQ   SEQUENCE   159 AA;  17201 MW;  364FE35A65E2C89D CRC64;


  Query Match             85.7%;  Score 18;  DB 1;  Length 159;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 eagxxs 6
             |||  |
Db     145 EAGSGS 150


Search completed: August 30, 2002, 15:11:55
Job time: 311 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 30, 2002, 15:05:49 ; Search time 41.29 Seconds
(without alignments)
25.139 Million cell updates/sec

Title: BASK-853-CLAIM4
Perfect score: 21
Sequence: 1 eagxxs 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	90.5	108	12	Q91TM3	Q91tm3 tupaia herp
2	19	90.5	122	17	Q9HPE3	Q9hpe3 halobacteri
3	19	90.5	130	15	O36890	O36890 human immun
4	19	90.5	131	6	Q9GK47	Q9gk47 galago cras
5	19	90.5	143	17	Q9YBX6	Q9ybx6 aeropyrum p
6	19	90.5	145	16	Q9KRU8	Q9kru8 vibrio chol
7	19	90.5	149	10	Q9FRJ5	Q9frj5 oryza sativ
8	19	90.5	151	10	Q9AWI9	Q9awi9 oryza sativ
9	19	90.5	152	10	Q9XIK5	Q9xik5 arabidopsis
10	19	90.5	161	4	Q9BS09	Q9bs09 homo sapien
11	19	90.5	161	4	Q9NWI1	Q9nwi1 homo sapien
12	19	90.5	164	2	Q54209	Q54209 streptomyce
13	19	90.5	170	6	Q9MZ11	Q9mz11 oryctolagus
14	19	90.5	192	10	Q94LT3	Q94lt3 oryza sativ
15	19	90.5	200	5	Q18144	Q18144 caenorhabdi
16	19	90.5	201	5	O62323	O62323 caenorhabdi
17	19	90.5	204	16	O53374	O53374 mycobacteri
18	19	90.5	212	16	Q92D00	Q92d00 listeria in
19	19	90.5	231	13	Q91306	Q91306 rana catesb
20	19	90.5	245	17	Q9HST1	Q9hst1 halobacteri
21	19	90.5	254	13	Q91307	Q91307 rana catesb
22	19	90.5	257	10	Q9XF64	Q9xf64 arabidopsis
23	19	90.5	257	10	Q9LZJ6	Q9lzj6 arabidopsis
24	19	90.5	260	16	Q99RZ3	Q99rz3 staphylococ
25	19	90.5	261	2	Q9ZIN7	Q9zin7 staphylococ
26	19	90.5	261	16	Q92NH7	Q92nh7 rhizobium m
27	19	90.5	266	12	Q88190	Q88190 soybean mos
28	19	90.5	267	12	Q88196	Q88196 soybean mos
29	19	90.5	269	11	Q9DCE4	Q9dce4 mus musculu
30	19	90.5	270	5	Q9VGK4	Q9vgk4 drosophila
31	19	90.5	285	2	Q9AIT5	Q9ait5 vibrio chol
32	19	90.5	287	13	O93503	O93503 xenopus lae
33	19	90.5	293	2	Q9L866	Q91866 azospirillu
34	19	90.5	297	13	Q91296	Q91296 rana catesb
35	19	90.5	302	16	Q9PCJ6	Q9pcj6 xylella fas
36	19	90.5	307	10	Q9FJG2	Q9fjg2 arabidopsis
37	19	90.5	311	13	Q90888	Q90888 gallus gall
38	19	90.5	311	13	Q90370	Q90370 coturnix co
39	19	90.5	313	13	Q9PUA6	Q9pua6 xenopus lae
40	19	90.5	323	4	Q9Y5Q3	Q9y5q3 homo sapien
41	19	90.5	323	4	Q9H1F1	Q9h1f1 homo sapien
42	19	90.5	333	13	Q98TS3	Q98ts3 brachydanio
43	19	90.5	346	10	Q9ATS6	Q9ats6 arundinella
44	19	90.5	350	16	Q981H9	Q981h9 rhizobium l
45	19	90.5	352	16	Q92WT2	Q92wt2 rhizobium m
46	19	90.5	356	13	Q98UK5	Q98uk5 brachydanio
47	19	90.5	356	13	O73679	O73679 brachydanio
48	19	90.5	357	10	O23101	O23101 arabidopsis
49	19	90.5	361	5	Q9BL91	Q9bl91 caenorhabdi
50	19	90.5	362	11	Q9JHQ1	Q9jhq1 rattus norv

ALIGNMENTS

RESULT 1
Q91TM3
ID Q91TM3 PRELIMINARY; PRT; 108 AA.
AC Q91TM3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE T73.
OS Tupaia herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RA Bahr U., Darai G.;
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree
RT Shrew) Herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Darai G., Bahr U.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF281817; AAK57118.1; -.
SQ SEQUENCE 108 AA; 11188 MW; 5732B2C61DBDE820 CRC64;

Query Match 90.5%; Score 19; DB 12; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 eagxxs 6

```
          |||  |
Db     30 EAGASS 35


RESULT   2
Q9HPE3
ID   Q9HPE3      PRELIMINARY;      PRT;   122 AA.
AC   Q9HPE3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   VNG1678H.
GN   VNG1678H.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005076; AAG19927.1; -.
KW   Complete proteome.
SQ   SEQUENCE   122 AA;  12001 MW;  BE2432416C2BA256 CRC64;


  Query Match             90.5%;  Score 19;  DB 17;  Length 122;
  Best Local Similarity   66.7%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     57 EAGASS 62


RESULT   3
O36890
ID   O36890      PRELIMINARY;      PRT;   130 AA.
AC   O36890;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GAG POLYPROTEIN (FRAGMENT).
GN   GAG.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97445059; PubMed=9300048;
RA   Leigh Brown A.J., Lobidel D., Wade C.M., Rebus S., Phillips N.,
RA   Brettle R.P., France A.J., Leen C.S., McMenamin J., McMillan A.,
RA   Maw R.D., Mulcahy F., Robertson J.R., Sankar K.N., Scott G., Wyld R.,
RA   Peutherer J.F.;
RT   "The molecular epidemiology of human immunodeficiency virus type 1 in
RT   six cities in Britain and Ireland.";
RL   Virology 235:166-177(1997).
DR   EMBL; AF014287; AAC58368.1; -.
DR   InterPro; IPR000071; Retroviral_gag_p17.
DR   Pfam; PF00540; gag_p17; 1.
DR   PRINTS; PR00234; HIV1MATRIX.
KW   AIDS; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   NON_TER     130    130
SQ   SEQUENCE   130 AA;  14476 MW;  9053DD2EF3A3E00F CRC64;


  Query Match             90.5%;  Score 19;  DB 15;  Length 130;
  Best Local Similarity   66.7%;  Pred. No. 7.1e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     99 EAGSSS 104


RESULT   4
Q9GK47
ID   Q9GK47      PRELIMINARY;      PRT;   131 AA.
AC   Q9GK47;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RELAXIN-LIKE PROTEIN.
OS   Galago crassicaudatus (Thick-tailed galago) (Otolemur crassicaudatus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Otolemur.
OX   NCBI_TaxID=9463;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Klonisch T., Froehlich C., Tetens F., Fischer B., Hombach-Klonisch S.;
RT   "Molecular remodeling of members of the relaxin family during primate
RT   evolution.";
RL   Mol. Biol. Evol. 0:0-0(2001).
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; AF317624; AAG42317.1; -.
DR   InterPro; IPR000739; Insulin_IGF_relaxin.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
SQ   SEQUENCE   131 AA;  14414 MW;  F18AFA9ACFC85943 CRC64;


  Query Match             90.5%;  Score 19;  DB 6;  Length 131;
  Best Local Similarity   66.7%;  Pred. No. 7.1e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     54 EAGTSS 59


RESULT   5
Q9YBX6
ID   Q9YBX6      PRELIMINARY;      PRT;   143 AA.
AC   Q9YBX6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 15.7 KDA PROTEIN APE1474.
GN   APE1474.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
```

```
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
DR   EMBL; AP000061; BAA80472.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   143 AA;  15708 MW;  C40E29CBB0AF6892 CRC64;


  Query Match             90.5%;  Score 19;  DB 17;  Length 143;
  Best Local Similarity   66.7%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     85 EAGAAS 90


RESULT   6
Q9KRU8
ID   Q9KRU8      PRELIMINARY;      PRT;   145 AA.
AC   Q9KRU8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN VC1536.
GN   VC1536.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004231; AAF94690.1; -.
DR   TIGR; VC1536; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   145 AA;  15726 MW;  976E1F5EB50DB0EC CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 145;
  Best Local Similarity   66.7%;  Pred. No. 7.9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    114 EAGSTS 119


RESULT   7
Q9FRJ5
ID   Q9FRJ5      PRELIMINARY;      PRT;   149 AA.
AC   Q9FRJ5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 15.3 KDA PROTEIN.
GN   OSJNBB0064P21.9.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA   Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA   Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA   Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBb0064P21 genomic sequence.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC073166; AAG46108.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   149 AA;  15284 MW;  8B71E92310872766 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 149;
  Best Local Similarity   66.7%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    111 EAGASS 116


RESULT   8
Q9AWI9
ID   Q9AWI9      PRELIMINARY;      PRT;   151 AA.
AC   Q9AWI9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   P0489A05.8 PROTEIN.
GN   P0489A05.8.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0489A05.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003105; BAB32988.1; -.
SQ   SEQUENCE   151 AA;  16103 MW;  6E942A203BC62C11 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 151;
  Best Local Similarity   66.7%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    106 EAGAAS 111


RESULT   9
Q9XIK5
ID   Q9XIK5      PRELIMINARY;      PRT;   152 AA.
AC   Q9XIK5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   T10O24.6.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Brooks S., Buehler E., Chao Q., Dunn P., Khan S., Kim C.,
RA   Walker M., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
RA   Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,
```

```
RA   Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
RA   Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RT   "Genomic sequence for Arabidopsis thaliana BAC T10O24 from Chromosome
RT   1.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007067; AAD39566.1; -.
SQ   SEQUENCE   152 AA;  17676 MW;  A7053F4DA73C3490 CRC64;


  Query Match            90.5%;  Score 19;  DB 10;  Length 152;
  Best Local Similarity   66.7%;  Pred. No. 8.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    138 EAGTTS 143


RESULT  10
Q9BS09
ID   Q9BS09      PRELIMINARY;      PRT;   161 AA.
AC   Q9BS09;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 17.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON ADENOCARCINOMA;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005805; AAH05805.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   161 AA;  17058 MW;  E4098AB1F0A5D706 CRC64;


  Query Match            90.5%;  Score 19;  DB 4;  Length 161;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     43 EAGSSS 48


RESULT  11
Q9NWI1
ID   Q9NWI1      PRELIMINARY;      PRT;   161 AA.
AC   Q9NWI1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CDNA FLJ20847 FIS, CLONE ADKA01746.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ADIPOSE TISSUE;
RA   Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA   Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA   Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK000854; BAA91399.1; -.
SQ   SEQUENCE   161 AA;  17118 MW;  11098AB1EA15D71C CRC64;


  Query Match            90.5%;  Score 19;  DB 4;  Length 161;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     43 EAGSSS 48


RESULT  12
Q54209
ID   Q54209      PRELIMINARY;      PRT;   164 AA.
AC   Q54209;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   FABD, FABH, FABC, FABB, AND ORF5 GENES.
OS   Streptomyces glaucescens.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1907;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GLA.0;
RX   MEDLINE=95352622; PubMed=7626609;
RA   Summers R.G., Ali A., Shen B., Wessel W.A., Hutchinson C.R.;
RT   "Malonyl-coenzyme A:acyl carrier protein acyltransferase of
RT   Streptomyces glaucescens: a possible link between fatty acid and
RT   polyketide biosynthesis.";
RL   Biochemistry 34:9389-9402(1995).
DR   EMBL; L43074; AAA99450.1; -.
SQ   SEQUENCE   164 AA;  18203 MW;  CB0ECF031044BB09 CRC64;


  Query Match            90.5%;  Score 19;  DB 2;  Length 164;
  Best Local Similarity   66.7%;  Pred. No. 9e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     55 EAGTAS 60


RESULT  13
Q9MZ11
ID   Q9MZ11      PRELIMINARY;      PRT;   170 AA.
AC   Q9MZ11;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   5-HT1A (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RETINA;
RA   Pootanakit K., Hunter D.D., Brunken W.J.;
RT   "5-HT1A and 5-HT7 Receptor Expression in the Mammalian Retina.";
RL   Brain Res. 0:0-0(2000).
DR   EMBL; AF269231; AAF76184.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
FT   NON_TER       1      1
FT   NON_TER     170    170
SQ   SEQUENCE   170 AA;  18518 MW;  42A5B4CF917B3250 CRC64;


  Query Match            90.5%;  Score 19;  DB 6;  Length 170;
  Best Local Similarity   66.7%;  Pred. No. 9.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy      1 eagxxs 6
          |||  |
Db    142 EAGAAS 147


RESULT  14
Q94LT3
ID   Q94LT3      PRELIMINARY;      PRT;   192 AA.
AC   Q94LT3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 19.8 KDA PROTEIN.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA   Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA   Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA   Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA   White O., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBb0011A08 genomic sequence.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC034258; AAK54288.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   192 AA;  19819 MW;  4CE8C88AE83DF374 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 192;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     40 EAGAAS 45


RESULT  15
Q18144
ID   Q18144      PRELIMINARY;      PRT;   200 AA.
AC   Q18144;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 22.0 KDA PROTEIN.
GN   C25A8.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Latreille P., Stellyes L.;
RT   "The sequence of C. elegans cosmid C25A8.";
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U61958; AAB03180.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   200 AA;  22012 MW;  66A23EDA709C66B2 CRC64;


  Query Match             90.5%;  Score 19;  DB 5;  Length 200;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    109 EAGAAS 114


RESULT  16
O62323
ID   O62323      PRELIMINARY;      PRT;   201 AA.
AC   O62323;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   R02D5.7 PROTEIN.
GN   R02D5.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Matthews L.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z78015; CAB01436.1; -.
SQ   SEQUENCE   201 AA;  22266 MW;  EC0423A8D7DDE4FE CRC64;


  Query Match             90.5%;  Score 19;  DB 5;  Length 201;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    110 EAGAAS 115


RESULT  17
O53374
ID   O53374      PRELIMINARY;      PRT;   204 AA.
AC   O53374;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 21.6 KDA PROTEIN.
GN   MOAE3 OR RV3322C OR MTV016.22C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
```

```
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
DR   EMBL; AL021841; CAA17094.1; -.
DR   TubercuList; Rv3322c; -.
DR   InterPro; IPR000051; SAM_bind.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   204 AA;  21614 MW;  13C5CB74C9C4B07F CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 204;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    153 EAGTAS 158


RESULT  18
Q92D00
ID   Q92D00      PRELIMINARY;      PRT;   212 AA.
AC   Q92D00;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIN1021 PROTEIN.
GN   LIN1021.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596167; CAC96252.1; -.
DR   ListiList; LIN01021; -.
KW   Complete proteome.
SQ   SEQUENCE   212 AA;  23672 MW;  CB73DB2965A08F99 CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 212;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     95 EAGASS 100


RESULT  19
Q91306
ID   Q91306      PRELIMINARY;      PRT;   231 AA.
AC   Q91306;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MYOSIN VI (FRAGMENT).
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SACCULE;
RA   Solc C.F., Derfler B.H., Duyk G.M., Corey D.P.;
RT   "Molecular cloning of myosins from the bullfrog saccular macula: A
RT   candidate for the hair-cell adaptation motor.";
RL   Aud. Neurosci. 1:63-75(1994).
DR   EMBL; U14380; AAA65089.1; -.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
DR   ProDom; PD000355; myosin_head; 1.
FT   NON_TER       1      1
FT   NON_TER     231    231
SQ   SEQUENCE   231 AA;  25693 MW;  D3FFF5C343E6FAC8 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 231;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    124 EAGSTS 129


RESULT  20
Q9HST1
ID   Q9HST1      PRELIMINARY;      PRT;   245 AA.
AC   Q9HST1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   L-ISOASPARTYL PROTEIN CARBOXYL METHYLTRANSFERASE.
GN   PIMT1 OR VNG0089G.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE004977; AAG18721.1; -.
DR   InterPro; IPR000682; PCMT.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF01135; PCMT; 1.
KW   Transferase; Methyltransferase; Complete proteome.
SQ   SEQUENCE   245 AA;  26216 MW;  A26FBBBCFAA5DB78 CRC64;


  Query Match             90.5%;  Score 19;  DB 17;  Length 245;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
```

```
          |||  |
Db    151 EAGAAS 156


RESULT  21
Q91307
ID   Q91307      PRELIMINARY;      PRT;   254 AA.
AC   Q91307;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MYOSIN VI (FRAGMENT).
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SACCULE;
RA   Solc C.F., Derfler B.H., Duyk G.M., Corey D.P.;
RT   "Molecular cloning of myosins from the bullfrog saccular macula: A
RT   candidate for the hair-cell adaptation motor.";
RL   Aud. Neurosci. 1:63-75(1994).
DR   EMBL; U14381; AAA65090.1; -.
DR   HSSP; P10587; 1BR2.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 3.
DR   ProDom; PD000355; myosin_head; 1.
FT   NON_TER       1      1
FT   NON_TER     254    254
SQ   SEQUENCE   254 AA;  29039 MW;  DB8395E86BD6DE93 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 254;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    169 EAGSTS 174


RESULT  22
Q9XF64
ID   Q9XF64      PRELIMINARY;      PRT;   257 AA.
AC   Q9XF64;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RING-H2 ZINC FINGER PROTEIN ATL5.
GN   ATL5.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99408259; PubMed=10480382;
RA   Salinas-Mondragon R.E., Garciduenas-Pina C., Guzman P.;
RT   "Early elicitor induction in members of a novel multigene family
RT   coding for highly related RING-H2 proteins in Arabidopsis thaliana.";
RL   Plant Mol. Biol. 40:579-590(1999).
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR   EMBL; AF132015; AAD33583.1; -.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
KW   Zinc-finger.
SQ   SEQUENCE   257 AA;  28608 MW;  07BCEF8CEC928C96 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 257;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    184 EAGSSS 189


RESULT  23
Q9LZJ6
ID   Q9LZJ6      PRELIMINARY;      PRT;   257 AA.
AC   Q9LZJ6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RING-H2 ZINC FINGER PROTEIN ATL5.
GN   F26K9_120.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bloecker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR   EMBL; AL162651; CAB83119.1; -.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
KW   Zinc-finger.
SQ   SEQUENCE   257 AA;  28592 MW;  B6B7595DFF528431 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 257;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    184 EAGSSS 189


RESULT  24
Q99RZ3
ID   Q99RZ3      PRELIMINARY;      PRT;   260 AA.
AC   Q99RZ3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MODA PROTEIN.
GN   MODA OR SA2074.
OS   Staphylococcus aureus (strain N315).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=158879;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA   Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA   Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA   Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA   Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA   Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
```

```
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003136; BAB43371.1; -.
DR   HSSP; P37329; 1WOD.
KW   Complete proteome.
SQ   SEQUENCE   260 AA;  29117 MW;  7A5D4A01A4482C4D CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 260;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    227 EAGATS 232


RESULT  25
Q9ZIN7
ID   Q9ZIN7      PRELIMINARY;      PRT;  261 AA.
AC   Q9ZIN7;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MODA.
GN   MODA.
OS   Staphylococcus carnosus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1281;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TM300;
RX   MEDLINE=98340502; PubMed=9675851;
RA   Neubauer H., Pantel I., Gotz F.;
RT   "Characterization of moeB--part of the molybdenum cofactor
RT   biosynthesis gene cluster in Staphylococcus carnosus.";
RL   FEMS Microbiol. Lett. 164:55-62(1998).
DR   EMBL; AF109295; AAC83133.1; -.
DR   HSSP; P37329; 1WOD.
SQ   SEQUENCE   261 AA;  29203 MW;  126A2D314BBAFB13 CRC64;


  Query Match             90.5%;  Score 19;  DB 2;  Length 261;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    228 EAGATS 233


RESULT  26
Q92NH7
ID   Q92NH7      PRELIMINARY;      PRT;  261 AA.
AC   Q92NH7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
GN   SMC01615.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21368234; PubMed=11474104;
RA   Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA   Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA   Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA   Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA   Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA   Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA   Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA   Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA   Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA   Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT   "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL   Science 293:668-672(2001).
DR   EMBL; AL591790; CAC46806.1; -.
KW   Complete proteome.
SQ   SEQUENCE   261 AA;  28115 MW;  031D03708E1084CB CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 261;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     16 EAGTAS 21


RESULT  27
Q88190
ID   Q88190      PRELIMINARY;      PRT;  266 AA.
AC   Q88190;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COAT PROTEIN.
OS   Soybean mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=12222;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SA;
RA   Chu R., Leng X., Bao Y., Pu Z., Pan N., Chen Z.;
RT   "Amplification of soybean mosaic virus coat protein gene by polymerase
RT   chain reaction and its sequence analysis.";
RL   Acta Bot. Sin. 34:523-528(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SA;
RA   Xu L.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U25673; AAA70095.1; -.
DR   InterPro; IPR001592; Poty_coat.
DR   Pfam; PF00767; Poty_coat; 1.
SQ   SEQUENCE   266 AA;  30084 MW;  4E08AFE7D434307F CRC64;


  Query Match             90.5%;  Score 19;  DB 12;  Length 266;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     25 EAGTSS 30


RESULT  28
Q88196
ID   Q88196      PRELIMINARY;      PRT;  267 AA.
AC   Q88196;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COAT PROTEIN (FRAGMENT).
OS   Soybean mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=12222;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHINESE;
RA   Chu R.;
RT   "CDNA sequence of the gene encoding coat protein of SMV.";
RL   Nucleic Acids Res. 0:0-0(0).
DR   EMBL; X63771; CAA45307.1; -.
DR   InterPro; IPR001592; Poty_coat.
DR   Pfam; PF00767; Poty_coat; 1.
FT   NON_TER      1      1
SQ   SEQUENCE   267 AA;  30104 MW;  220E42F2595BE059 CRC64;


  Query Match             90.5%;  Score 19;  DB 12;  Length 267;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     26 EAGTSS 31


RESULT  29
Q9DCE4
ID   Q9DCE4      PRELIMINARY;      PRT;   269 AA.
AC   Q9DCE4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   0610039P13RIK PROTEIN.
GN   0610039P13RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK002854; BAB22409.1; -.
DR   MGD; MGI:1921346; 0610039P13Rik.
DR   InterPro; IPR000636; Cation_chan_non_lig.
SQ   SEQUENCE   269 AA;  31242 MW;  B549CB553DEB6568 CRC64;


  Query Match             90.5%;  Score 19;  DB 11;  Length 269;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     71 EAGSAS 76


RESULT  30
Q9VGK4
ID   Q9VGK4      PRELIMINARY;      PRT;   270 AA.
AC   Q9VGK4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG14714 PROTEIN.
GN   CG14714.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003692; AAF54673.1; -.
DR   HSSP; P28827; 1RPM.
DR   FlyBase; FBgn0037929; CG14714.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_prot_phphtase.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   270 AA;  30259 MW;  5C136F3135CAD001 CRC64;


  Query Match             90.5%;  Score 19;  DB 5;  Length 270;
```

```
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    169 EAGSTS 174


RESULT  31
Q9AIT5
ID   Q9AIT5      PRELIMINARY;      PRT;   285 AA.
AC   Q9AIT5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ICMF-LIKE PROTEIN (FRAGMENT).
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=569B;
RX   MEDLINE=20434574; PubMed=10981695;
RA   Das S., Chakrabortty A., Banerjee R., Roychoudhury S., Chaudhuri K.;
RT   "Comparison of global transcription responses allows identification of
RT   Vibrio cholerae genes differentially expressed following infection.";
RL   FEMS Microbiol. Lett. 190:87-91(2000).
DR   EMBL; AF239737; AAK27321.1; -.
FT   NON_TER       1      1
FT   NON_TER     285    285
SQ   SEQUENCE   285 AA;  32165 MW;  BDDE7FA9021F5661 CRC64;


  Query Match             90.5%;  Score 19;  DB 2;  Length 285;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    210 EAGSAS 215


RESULT  32
O93503
ID   O93503      PRELIMINARY;      PRT;   287 AA.
AC   O93503;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE.
GN   MARCKS.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98030614; PubMed=9361009;
RA   Shi Y., Sullivan S.K., Pitterle D.M., Kennington E.A., Graff J.M.,
RA   Blackshear P.J.;
RT   "Mechanisms of MARCKS gene activation during Xenopus development.";
RL   J. Biol. Chem. 272:29290-29300(1997).
DR   EMBL; AF017299; AAC61897.1; -.
DR   InterPro; IPR002101; MARCKS.
DR   Pfam; PF02063; MARCKS; 1.
DR   PRINTS; PR00963; MARCKS.
DR   PROSITE; PS00826; MARCKS_1; 1.
DR   PROSITE; PS00827; MARCKS_2; 1.
KW   Kinase.
SQ   SEQUENCE   287 AA;  29147 MW;  35CB7AE6090ED3C1 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 287;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    110 EAGSTS 115


RESULT  33
Q9L866
ID   Q9L866      PRELIMINARY;      PRT;   293 AA.
AC   Q9L866;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 32.1 KDA PROTEIN.
OS   Azospirillum brasilense.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC   Azospirillum.
OX   NCBI_TaxID=192;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ma L., Zhao Y., Wang J., Li J.;
RT   "Sequence and function analysis of draTG genes downstream ORFs from
RT   Azospirillum brasilense Yu62.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF216815; AAF61911.1; -.
DR   InterPro; IPR003310; DNA_glycosylase.
KW   Hypothetical protein.
SQ   SEQUENCE   293 AA;  32063 MW;  37417EA008F6BD61 CRC64;


  Query Match             90.5%;  Score 19;  DB 2;  Length 293;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    168 EAGAAS 173


RESULT  34
Q91296
ID   Q91296      PRELIMINARY;      PRT;   297 AA.
AC   Q91296;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MYOSIN VI (FRAGMENT).
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SACCULE;
RA   Solc C.F., Derfler B.H., Duyk G.M., Corey D.P.;
RT   "Molecular cloning of myosins from the bullfrog saccular macula: A
RT   candidate for the hair-cell adaptation motor.";
RL   Aud. Neurosci. 1:63-75(1994).
DR   EMBL; U14370; AAA65079.1; -.
DR   HSSP; P10587; 1BR2.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
DR   ProDom; PD000355; myosin_head; 1.
FT   NON_TER       1      1
FT   NON_TER     297    297
SQ   SEQUENCE   297 AA;  33755 MW;  46EE6C78A8ED530D CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 297;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
```

```
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    169 EAGSTS 174


RESULT  35
Q9PCJ6
ID   Q9PCJ6      PRELIMINARY;      PRT;   302 AA.
AC   Q9PCJ6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN XF1783.
GN   XF1783.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE004000; AAF84591.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   302 AA;  32047 MW;  758CC61DE4BE3590 CRC64;


  Query Match            90.5%;  Score 19;  DB 16;  Length 302;
  Best Local Similarity  66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    106 EAGTAS 111


RESULT  36
Q9FJG2
ID   Q9FJG2      PRELIMINARY;      PRT;   307 AA.
AC   Q9FJG2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   GB|AAD26962.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryotâ; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=99087489; PubMed=9872454;
RA   Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT   Sequence features of the regions of 1,013,767 bp covered by sixteen
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:297-308(1998).
DR   EMBL; AB015473; BAB08399.1; -.
SQ   SEQUENCE   307 AA;  35727 MW;  23CDB6C127CE90D1 CRC64;


  Query Match            90.5%;  Score 19;  DB 10;  Length 307;
  Best Local Similarity  66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    207 EAGTSS 212


RESULT  37
Q90888
ID   Q90888      PRELIMINARY;      PRT;   311 AA.
AC   Q90888;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MAFB.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95021288; PubMed=7935473;
RA   Kataoka K., Fujiwara K.T., Noda M., Nishizawa M.;
RT   "MafB,a new Maf family transcription activator that can associate with
RT   Maf and Fos but not with Jun.";
RL   Mol. Cell. Biol. 14:7581-7591(1994).
DR   EMBL; D28600; BAA05938.1; -.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   311 AA;  35467 MW;  DDAE7F698B7D3ABA CRC64;


  Query Match            90.5%;  Score 19;  DB 13;  Length 311;
  Best Local Similarity  66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    296 EAGSTS 301


RESULT  38
Q90370
ID   Q90370      PRELIMINARY;      PRT;   311 AA.
AC   Q90370;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MAFB PROTEIN.
```

```
GN   MAFB.
OS   Coturnix coturnix japonica (Japanese quail).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Coturnix.
OX   NCBI_TaxID=93934;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96180718; PubMed=8620536;
RA   Sieweke M.H., Tekotte H., Frampton J., Graf T.;
RT   "MafB is an interaction partner and repressor of Ets-1 that inhibits
RT   erythroid differentiation.";
RL   Cell 85:49-60(1996).
DR   EMBL; X96511; CAA65360.1; -.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   311 AA;  35476 MW;  7D1F3FA05D5CD683 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 311;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    296 EAGSTS 301


RESULT  39
Q9PUA6
ID   Q9PUA6      PRELIMINARY;      PRT;   313 AA.
AC   Q9PUA6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BZIP TRANSCRIPTION FACTOR MAFB.
GN   MAFB.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21152895; PubMed=11231068;
RA   Ishibashi S., Yasuda K.;
RT   "Distinct roles of maf genes during Xenopus lens development.";
RL   Mech. Dev. 101:155-166(2001).
DR   EMBL; AF202058; AAF08316.1; -.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   313 AA;  35714 MW;  8E697A00A928BF95 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 313;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    298 EAGSTS 303


RESULT  40
Q9Y5Q3
ID   Q9Y5Q3      PRELIMINARY;      PRT;   323 AA.
AC   Q9Y5Q3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MAFB/KREISLER BASIC REGION/LEUCINE ZIPPER TRANSCRIPTION FACTOR.
GN   MAFB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BONE MARROW;
RX   MEDLINE=99375320; PubMed=10444328;
RA   Wang P.W., Eisenbart J.D., Cordes S.P., Barsh G.S., Stoffel M.,
RA   Le Beau M.M.;
RT   "Human KRML (MAFB): cDNA cloning, genomic structure, and evaluation as
RT   a candidate tumor suppressor gene in myeloid leukemias.";
RL   Genomics 59:275-281(1999).
DR   EMBL; AF134157; AAD30106.1; -.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   323 AA;  35829 MW;  AE4DC23408E36E55 CRC64;


  Query Match             90.5%;  Score 19;  DB 4;  Length 323;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    308 EAGSTS 313


RESULT  41
Q9H1F1
ID   Q9H1F1      PRELIMINARY;      PRT;   323 AA.
AC   Q9H1F1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DJ644L1.1 (KREISLER (MOUSE) MAF-RELATED LEUCINE ZIPPER HOMOLOG).
GN   KRML.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ramsay H.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035665; CAB75863.1; -.
DR   InterPro; IPR001871; bZIP.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   323 AA;  35792 MW;  A0F3C09F8936CB16 CRC64;


  Query Match             90.5%;  Score 19;  DB 4;  Length 323;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    308 EAGSTS 313


RESULT  42
Q98TS3
ID   Q98TS3      PRELIMINARY;      PRT;   333 AA.
AC   Q98TS3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HSP70 BINDING PROTEIN.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
```

```
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Guerriero V., Raynes D.A.;
RT   "Hsp70 binding protein from zebra fish (HspBPF).";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY024336; AAG61257.1; -.
DR   InterPro; IPR000225; Armadillo.
DR   Pfam; PF00514; Armadillo_seg; 2.
DR   SMART; SM00185; ARM; 2.
SQ   SEQUENCE   333 AA;  37269 MW;  E7C5ABD12F41D23E CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 333;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     23 EAGSAS 28


RESULT  43
Q9ATS6
ID   Q9ATS6      PRELIMINARY;      PRT;   346 AA.
AC   Q9ATS6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   TEOSINTE BRANCHED1 PROTEIN (FRAGMENT).
GN   TB1.
OS   Arundinella hirta.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Arundinelleae; Arundinella.
OX   NCBI_TaxID=79825;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21165336; PubMed=11264415;
RA   Lukens L., Doebley J.;
RT   "Molecular evolution of the teosinte branched gene among maize and
RT   related rrasses.";
RL   Mol. Biol. Evol. 18:627-638(2001).
DR   EMBL; AF322131; AAK37493.1; -.
FT   NON_TER       1      1
FT   NON_TER     346    346
SQ   SEQUENCE   346 AA;  36902 MW;  BFB5F29CD7449C89 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 346;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    280 EAGAAS 285


RESULT  44
Q981H9
ID   Q981H9      PRELIMINARY;      PRT;   350 AA.
AC   Q981H9;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   MLR9369 PROTEIN.
GN   MLR9369.
OS   Rhizobium loti (Mesorhizobium loti).
OG   Plasmid pMLa.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003016; BAB54976.1; -.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   350 AA;  39172 MW;  BB075F9345BB9362 CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 350;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    108 EAGASS 113


RESULT  45
Q92WT2
ID   Q92WT2      PRELIMINARY;      PRT;   352 AA.
AC   Q92WT2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ADENYLATE CYCLASE PROTEIN.
GN   SMB20257.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymB (megaplasmid 2).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396508; PubMed=11481431;
RA   Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA   Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA   Golding B., Puehler A.;
RT   "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT   fixing endosymbiont Sinorhizobium meliloti.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR   EMBL; AL603642; CAC48647.1; -.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   352 AA;  37018 MW;  F4AE6710196E06EF CRC64;


  Query Match             90.5%;  Score 19;  DB 16;  Length 352;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     13 EAGTSS 18


RESULT  46
Q98UK5
ID   Q98UK5      PRELIMINARY;      PRT;   356 AA.
AC   Q98UK5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TRANSCRIPTION FACTOR MAFB.
GN   MAFB.
```

```
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21064923; PubMed=11134968;
RA   Kajihara M., Kawauchi S., Kobayashi M., Ogino H., Takahashi S.,
RA   Yasuda K.;
RT   "Isolation, Characterization, and Expression Analysis of Zebrafish
RT   Large Mafs.";
RL   J. Biochem. 129:139-146(2001).
DR   EMBL; AB006322; BAB21102.1; -.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   356 AA;  40233 MW;  DE4C96B62C058865 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 356;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    341 EAGSTS 346


RESULT  47
O73679
ID   O73679      PRELIMINARY;      PRT;   356 AA.
AC   O73679;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TRANSCRIPTION FACTOR VAL.
GN   VAL OR VALENTINO.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98165393; PubMed=9425134;
RA   Moens C.B., Cordes S.P., Giorgianni M.W., Barsh G.S., Kimmel C.B.;
RT   "Equivalence in the genetic control of hindbrain segmentation in fish
RT   and mouse.";
RL   Development 125:381-391(1998).
DR   EMBL; AF006641; AAC18821.1; -.
DR   ZFIN; ZDB-GENE-980526-515; val.
DR   InterPro; IPR001871; bZIP.
DR   Pfam; PF03131; bZIP_Maf; 1.
DR   SMART; SM00338; BRLZ; 1.
SQ   SEQUENCE   356 AA;  40243 MW;  07420DB0F6CD08F1 CRC64;


  Query Match             90.5%;  Score 19;  DB 13;  Length 356;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    341 EAGSTS 346


RESULT  48
O23101
ID   O23101      PRELIMINARY;      PRT;   357 AA.
AC   O23101;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   A_TM018A10.10 PROTEIN.
GN   A_TM018A10.10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Dempsey S., Harper M.;
RT   "The sequence of A. thaliana TM018A10.";
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   WashU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF013294; AAB62869.1; -.
DR   InterPro; IPR001810; F-box.
DR   Pfam; PF00646; F-box; 1.
DR   SMART; SM00256; FBOX; 1.
SQ   SEQUENCE   357 AA;  40078 MW;  B1683A07BF630633 CRC64;


  Query Match             90.5%;  Score 19;  DB 10;  Length 357;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     91 EAGSSS 96


RESULT  49
Q9BL91
ID   Q9BL91      PRELIMINARY;      PRT;   361 AA.
AC   Q9BL91;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 40.2 KDA PROTEIN.
GN   Y18H1A.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Bemis G., Lamar B., Courtney L., Wohldmann P., Harrison M.;
RT   "The sequence of C. elegans cosmid Y18H1A.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
```

```
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024751; AAK21510.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   361 AA;  40201 MW;  98F44C3F87D59625 CRC64;


  Query Match             90.5%;  Score 19;  DB 5;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db     80 EAGSSS 85


RESULT  50
Q9JHQ1
ID   Q9JHQ1     PRELIMINARY;      PRT;   362 AA.
AC   Q9JHQ1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TITIN (FRAGMENT).
GN   TTN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKELETAL MUSCLE;
RX   MEDLINE=20490785; PubMed=11034912;
RA   Person V., Kostin S., Suzuki K., Labeit S., Schaper J.;
RT   "Antisense oligonucleotide experiments elucidate the essential role of
RT   titin in sarcomerogenesis in adult rat cardiomyocytes in long-term
RT   culture.";
RL   J. Cell Sci. 113:3851-3859(2000).
DR   EMBL; AJ401157; CAB95001.1; -.
DR   HSSP; P56276; 1TLK.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
FT   NON_TER     362    362
SQ   SEQUENCE   362 AA;  39601 MW;  E8E6CE65BB7F4ED8 CRC64;


  Query Match             90.5%;  Score 19;  DB 11;  Length 362;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 eagxxs 6
          |||  |
Db    317 EAGSSS 322


Search completed: August 30, 2002, 15:11:35
Job time: 346 sec
```